AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

2024, Vol. 153, No. 2, 339–351
https://doi.org/10.1037/xge0001496

Temporal Crowding With Central Vision Reveals the Fragility of Visual Representations

Tomer Sahar[1, 2] and Yaffa Yeshurun[1]
[1] School of Psychological Sciences, The Institute of Information Processing and Decision Making, University of Haifa
[2] Department of Psychology, The Open University of Israel

This study examined whether temporal crowding—the impaired object identification when distracting objects precede and succeed it—occurs at the fovea and if so whether its magnitude is reduced. We presented a central sequence of three oriented items separated by relatively long intervals (200/400 ms) and used an orientation estimation task with mixture-model analyses. We found clear evidence of temporal crowding with central vision, even with 400 ms intervals. Critically, reduced encoding precision surfaced as a robust and unique characteristic of temporal crowding. The magnitude of central and peripheral temporal crowding was similar suggesting the involvement of higher visual areas. Precision impairment emerged even when only the target contained orientation information, excluding "response competition" as the sole interference mechanism; yet it was larger when all items included orientation information, underscoring the importance of orientation-selective mechanisms. Overall, we show that even with central simple stimuli, the formation of a stable visual representation is surprisingly slow.

Public Significance Statement

Our study qualifies the common view that the visual system generates internal representations of visual objects very fast. We show that the internal representation of a simple oriented line is less precise when other orientated lines precede and succeed it, even when these stimuli are separated by long intervals and presented to the center of the visual field, where visual information is processed with the highest acuity. This suggests that generating a stable visual representation requires a surprisingly long time. Importantly, the finding that a similar interference emerges across the visual field suggests that the mechanisms underlying this long-lasting temporal interference involve processes taking place at later, higher visual areas. Additionally, this study elaborates on the conditions under which this long-lasting temporal interference may occur, thereby helping those who wish to avoid it.

Keywords: crowding, estimation task, mixture model, fovea, precision

There are several visual phenomena that demonstrate the imperfection of our visual system. These limitations are not just arbitrary failures, but consistent and systematic. For example, an object in our visual field is harder to identify when other items are also present than when it appears in isolation. For example, if the other items (the distractors) are presented at the same time as the to-be-identified object (the target), only at different locations, the impaired identification is attributed to "spatial crowding." When items are presented to the same location as the target, only before and after its presentation, the impaired identification is attributed to "temporal crowding." But what is the exact nature of these limitations? And importantly, why do they manifest? We know quite a lot about spatial crowding (recent reviews: Manassi & Whitney, 2018; Strasburger, 2020), but relatively little about temporal crowding, particularly in its most "pure" temporal form—when there is no spatial interference, that is, only a single item is presented at each time point; all stimuli are presented sequentially to the same location at different points in time (Bonneh et al., 2007; Tkacz-Domb & Yeshurun, 2017, 2021; Yeshurun et al., 2015). The two types of crowding share some common characteristics. For example, in both spatial and

This article was published Online First November 13, 2023.
Tomer Sahar https://orcid.org/0000-0001-9156-5928
This research was funded by the Israel Science Foundation (ISF) Grant 1780/19 to Yaffa Yeshurun. The authors declare no conflicts of interest. The data material and code are available at the OSF repository (https://osf.io/2eacq/). Various subsets of the ideas and data of this study were presented in the following conferences: the European Conference on Visual Perception (ECVP, 2021, 2022), the annual meetings of the Vision Science Society (VSS, 2022), and the Psychonomic Society (Psychonomics, 2021). Additionally, a preprint was posted on PsyArXiv.

Tomer Sahar served as lead for data curation, software, validation, visualization, and writing–original draft, contributed equally to methodology and writing–review and editing, and served in a supporting role for conceptualization. Yaffa Yeshurun served as lead for conceptualization, funding acquisition, investigation, supervision and served in a supporting role for writing–original draft and writing–review and editing. Tomer Sahar and Yaffa Yeshurun contributed equally to formal analysis.

Correspondence concerning this article should be addressed to Tomer Sahar, Department of Psychology, The Open University of Israel, The Dorothy de Rothschild Campus, 1 University Road, P.O. Box 808, Ra'anana 43107, Israel. Email: tomelico@gmail.com

temporal crowding, the overall target identification is improved as its distance from the distractors is increased (in space or time). Yet, the two types of "crowding" also clearly differ from one another (Tkacz-Domb & Yeshurun, 2021). When considering the objects' location within the visual field, spatial crowding is mostly demonstrated at the periphery, and some researchers even state that spatial crowding is the most important limiting factor in peripheral vision (e.g., Pelli, 2008; Rosenholtz, 2016). Still, with some effort, it is possible to demonstrate spatial crowding at the fovea (e.g., Coates et al., 2018; Lev et al., 2014; Malania et al., 2007). What about temporal crowding? Does temporal crowding occur only with peripheral vision? Examining if and how temporal crowding changes across the visual field is not only important for our understanding of how visual representations are affected by temporal crowding but also to advance our understanding of the mechanisms underlying this temporal interference.

Bonneh et al. (2007) observed temporal crowding at the fovea only for strabismic amblyopes, not for normal observers. However, that study focused on amblyopia and was not designed to examine normal vision. All other studies of "pure" temporal crowding (i.e., without spatial interference) involved peripheral presentation. For instance, two studies (Tkacz-Domb & Yeshurun, 2017; Yeshurun et al., 2015) presented a sequence of three letters to the same peripheral location, one of which was the target. The stimulus onset asynchrony (SOA) between the letters varied across trials, with all SOAs longer than the limits of ordinary masking (i.e., longer than 100–150 ms; Breitmeyer, 1984; Breitmeyer & Ogmen, 2000, 2006; Enns, 2004; Enns & Di Lollo, 2000; Gorea, 1987). These studies revealed an impaired performance that lasts over long intervals (i.e., with SOAs longer than 400 ms). This long-lasting impairment was observed even when there was no temporal uncertainty, or when an attentional cue indicated the location of the sequence of letters.

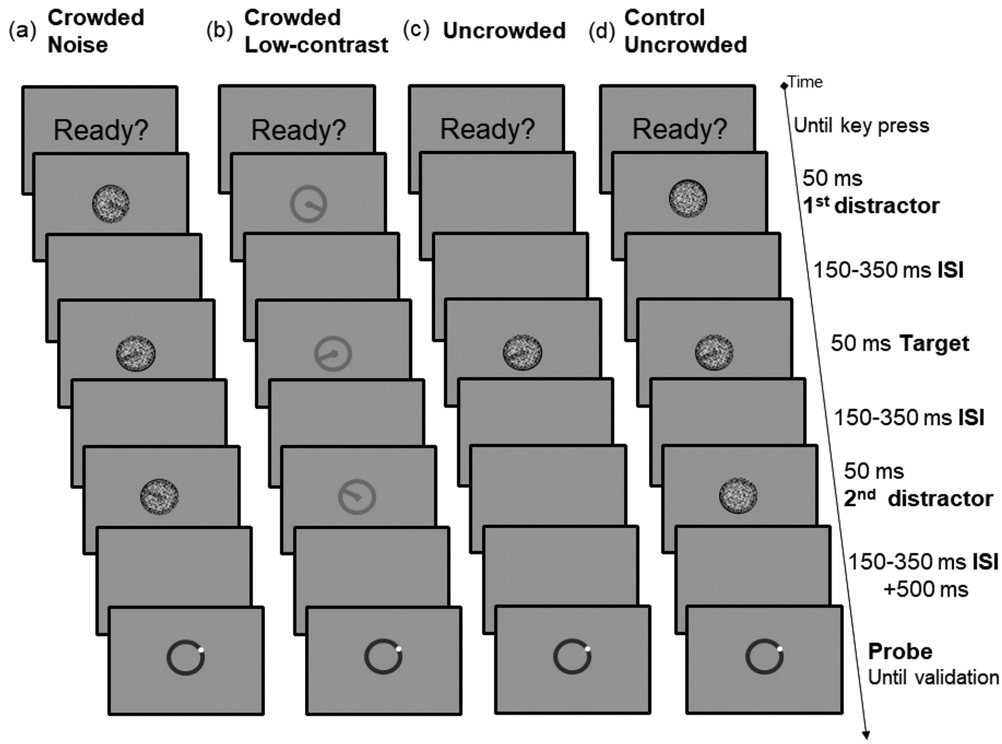
In a more recent study (Tkacz-Domb & Yeshurun, 2021), the sequence of letters was replaced with a sequence of oriented stimuli, the task was an orientation estimation task (i.e., continuous report), and the distribution of errors was analyzed using a mixture-modeling approach (detailed in the Method section), which allows classifying performance as a function of different error types. A similarly long-lasting temporal interference emerged with the orientation estimation task as was found with a discrete task (Tkacz-Domb & Yeshurun, 2017; Yeshurun et al., 2015), suggesting that temporal crowding does not depend on a specific task. Critically, this analysis showed that temporal crowding at the periphery: (a) reduced precision of the target encoding, which suggests that temporal crowding degrades the quality of target representation; (b) increased substitution errors (i.e., mistakenly reporting the orientation of a distractor instead of the target); (c) had no effect on the guess rate, suggesting that the signal-to-noise ratio (SNR) was not affected by temporal crowding. This pattern of results was observed regardless of stimuli duration or target-distractor similarity. In contrast, the opposite pattern was found in a study (Agaoglu et al., 2015) that used a similar orientation estimation task and mixture-modeling approach to examine masking (i.e., with short SOAs). Masking mainly affected the guess rate (SNR) but not the precision of the target encoding (quality of representation), suggesting that temporal crowding and ordinary masking are distinct phenomena.

Temporal crowding also qualitatively differs from spatial crowding. When Ester et al. (2014) examined the effect of spatial distance between items, with similar stimuli and a similar modeling approach, they found that reducing the spatial distance between items increased the guessing rate and the rate of substitution errors without affecting the precision of the target encoding. Whereas, as mentioned, temporal crowding (i.e., the effect of temporal distance between items) affected the precision of the target encoding and substitution rate but not the guess rate (Tkacz-Domb & Yeshurun, 2021). This different pattern of results shows that the two types of crowding likely reflect different processes. Given these differences between spatial and temporal crowding, we cannot tell whether temporal crowding is unique to the periphery based on studies of spatial crowding.

Thus, the current study was designed to examine whether temporal crowding can also be found with a foveal presentation, and if so whether its magnitude is reduced at the fovea, as was found with spatial crowding. Answering these questions is important for several reasons. First, examining how temporal crowding changes across the visual field offers a better understanding of the mechanisms involved. For instance, early processes of visual information are affected dramatically by the position within the visual field of the to-be-processed information, while later visual processes depend to a lesser degree on spatial position. Thus, the involvement of early processes in temporal crowding predicts that its magnitude would be reduced at the fovea while the involvement of higher processes does not predict such scaling. Second, the outcomes of this inquiry bear important theoretical implications regarding the rate of visual processing. Specifically, the fact that the distractors impair the precision of the target encoding even when separated by a very long SOA (475 ms; Tkacz-Domb & Yeshurun, 2021) suggests that a stable representation of the target is not consolidated even with such a long timescale. This arguably qualifies the common view that a robust representation is generated very fast—an initial representation may be generated fast, but it remains volatile for a long duration. Critically, the extent of this theoretical limitation depends on whether temporal crowding also occurs at the center of the visual field. If temporal crowding does not occur at the center (i.e., it is a uniquely peripheral phenomenon), then the surprisingly slow formation of a stable visual representation that is implied by temporal crowding would also be restricted to the periphery. If, however, temporal crowding does occur at the center, it would suggest that even a high-quality visual representation that is generated at the fovea can be degraded long after its formation. Finally, for those who wish to avoid the impairment brought about by temporal crowding, the study of how it varies across the visual field has critical methodological and applied implications. That is, if one is presenting a sequence of stimuli and wish to avoid interitem interference, for theoretical or applied reasons, it is critical to know the shortest interitem interval allowing to avoid such interference and if/how this interval varies across the visual field.

We conducted four experiments to examine temporal crowding at the fovea. In the crowded condition of all experiments, a sequence of three oriented items was presented to the center of the screen (Figure 1). The orientation of each item was determined randomly, and the task was to rotate a probe to reproduce the orientation of the target—the second item in the sequence. The SOA between the items was either 200 or 400 ms; it was constant within a trial and varied across trials. In the uncrowded condition only one item—the target—appeared. Because perceiving orientation is considerably easier at the fovea than the periphery (e.g., Virsu & Rovamo, 1979), the stimuli originally employed with peripheral presentation had to be modified to avoid ceiling effects (i.e., a too-easy

Figure 1
Stimuli and Design in Experiments 1–4



Note. A schematic (not scaled) illustration of the sequence of events in (a) the crowded trials of Experiments 1, 3, and 4; (b) the crowded trials of Experiment 2; (c) the uncrowded trials of Experiments 1, 3, and 4; (d) the control condition of Experiment 4. There were two possible target-distractor SOAs (200 and 400 ms). The SOA was constant within a trial but varied between trials. The task was to rotate the probe to match the target orientation. ISI = interstimulus interval; SOA = stimulus onset asynchrony.

task) due to foveal presentation. We used two different methods to make the task harder: the items in Experiments 1, 3, and 4 were embedded in random external noise, whereas in Experiment 2 we reduced the items' contrast. Because Experiments 1, 2, and 4 were conducted online, Experiment 3 was conducted as a lab validation under fully controlled conditions. The estimation errors were analyzed with the two-misreport mixture model (Jimenez et al., 2022; Kewan-Khalayly & Yashar, 2022; Shechter & Yashar, 2021; Yashar et al., 2019; see Mixture-Model Analysis section below). The main questions tested here were whether temporal crowding will emerge under foveal presentation (i.e., whether we will find a significant effect of SOA on some or all the parameters extracted with the mixture-modeling approach), and if so whether the pattern of results will be similar to that observed at the periphery. Finally, to better understand the processes underlying temporal crowding, Experiment 4 included a control condition in which the distractors did not include orientation information. It examined which factor is more determinantal, the mere presence of distractors or the fact that the distractors also include information at the task-relevant visual dimension (i.e., orientation). If temporal crowding merely reflects response competition generated by the orientations of the distractors, then we should not find the typical temporal interference in the control condition because its distractors do not include orientation information. Alternatively, if we find comparable temporal crowding in the control condition this will suggest that the processes mediating temporal crowding occur at regions that are indifferent to orientation.

Method

Participants

Based on Tkacz-Domb and Yeshurun (2021) and taking into consideration that online experiments may involve larger variability, we aimed for a sample size of at least 20 participants. A power test for analysis of variance (ANOVA) conducted with the R *pwr* package (Champely, 2020) using the effect size from Tkacz-Domb and Yeshurun (2021; Cohen's $f = .58$) and an alpha level of .05 confirmed that this sample size ($n = 20$) provides .95 power for detecting an effect. In the end, due to the exclusion procedures (details below), the final number of participants that were included in the statistical analysis was 21, 23, 20, and 23 in Experiments 1–4, respectively.

The participants were asked to provide their demographic information (age and gender) through a free-typing dialog box. They were asked to indicate their age by typing in the numbers, and their gender by typing: "M" for male, "F" for female, or "O" for other. The mean age in Experiments 1–4 was 25.1, 24.7, 26.6, and 23.9, respectively. Gender information—Experiment 1: 11 male, 10 female, 0 other; Experiment 2: 15 male, seven female, one other; Experiment 3: seven male, 13 female, 0 other; Experiment 4: eight male, 14 female, one other. The study was approved by the ethics committees of the University of Haifa (287/19) and the Open University of Israel. All experiments were performed in accordance with the guidelines and regulations of the Declaration of Helsinki.

Recruitment Procedure

For the online experiments (Experiments 1, 2, and 4), we recruited participants through *Prolific* (www.prolific.co). The participants were recruited if they were aged 18–35, native English speakers, and submitted at least 10 tasks previously, of which 90% were accepted. The participants' submission was accepted if they completed the entire task in a reasonable time, in full-screen mode, and without excessive repeated responses. The participants were paid 3.5£ for completing the task. For the lab experiment (Experiment 3), we recruited students from the Open University of Israel who were paid 30 Israeli Shekels for completing the experiment. All participants reported normal or corrected-to-normal vision and all signed a consent form. The study was approved by the ethics committees of the Open University of Israel and the University of Haifa (287/19).

Exclusion Procedure

Following Agaoglu et al. (2015), we calculated the overall performance score of each participant:

$$\text{Overall performance} = 1 - (\text{mean absolute error}/180). \quad (1)$$

If a participant always guessed the target's orientation, this score would be around .50, because the average of the absolute error will be around 90°. Only participants whose overall performance score was higher than .55 in all experimental conditions were included in the final statistical analysis. Thus, three, one, three, and one participants were excluded from Experiments 1–4, respectively.

Material

The task was implemented with PsychoPy (Peirce et al., 2019) and ran on *Pavlovia* (www.pavlovia.org). Mixture-model analyses and model comparisons were performed using the *Memtoolbox* package (Suchow et al., 2013). In the online experiments (Experiments 1, 2, and 4), we controlled the stimuli size using the "Virtual chin-rest" procedure (Li et al., 2020; Morys-Carter, 2021). This was not needed in the lab experiment (Experiment 3) in which we used a real chin rest, 56 cm from a 23.5″ LCD Eizo Foris monitor (1920 × 1080, 120 Hz) in a dimly lit room.

Stimuli

All stimuli appeared against a gray background (RGB = [128, 128, 128]). The initial "ready" display in each trial included the following text (in black, Arial font size 18): "Ready? Press the spacebar to continue" and the number of trials done out of the total trials. The various stimuli characteristics were chosen (and tested in a small online prestudy) to promote an overall performance that would be comparable in difficulty to that of Tkacz-Domb and Yeshurun (2021).

Experiment 1

The orientation items (target and distractors) were composed of a circle with a radius line (Figure 1a) that subtended 1.25 cm, generated using MS Paint. The width of the circle contour and the radius line was 1 pixel. In each trial of the crowded conditions, the orientation of the line within the circle was chosen randomly for each of the three items in the trial sequence with the constraint that their orientation will differ from each other by at least 30°. The items were embedded in a white noise patch. We generated three different noise patches using an online image generator (https://www.robson.plus/white-noise-image-generator/) that differed only in their noise pattern. Each patch had a radius of 1.25 cm, and we used Lunacy (https://icons8.com/lunacy) to overlay the patch on top of the circle and line; the noise opacity was set to 75%. For a given trial, one patch was chosen randomly and all three items of this trial were embedded in the selected patch. The probe was composed of a black semiopaque circle (RGB = [0, 0, 0], opacity 80%, radius 1.5 cm) and a white response dot (RGB = [255, 255, 255], 10 × 10 pixels) on its circumference. A feedback display was presented in the practice trials and it consisted of the probe with the response dot as indicated by the participant and a "feedback" dot that indicated the target's actual orientation. The feedback dot was green (RGB = [0, 255, 0]) if the absolute error was under 20° or red (RGB = [255, 0, 0]) if it was larger.

Experiment 2

The stimuli were identical to Experiment 1 except for the following: The orientation items were not embedded in noise (Figure 1b). Instead, their contrast was lowered by reducing the opacity of the circle and line to 11%.

Experiment 3

The stimuli were identical to Experiment 1 except for the following: The length of the radius line was increased to 1.75 cm to ensure overall performance is comparable (due to the fixed viewing distance of 56 cm and screen size).

Experiment 4

The stimuli were identical to Experiment 1 except for the following: The length of the radius of the circles and noise patch was increased to 1.5 cm. There was only a single noise patch type. In the control trials, the distractors' circles did not include the orientation line (Figure 1d).

Procedure

Experiments 1 and 2

Participants started by providing their age, and gender. They continued with the "virtual chin-rest" procedure and then went through detailed instructions displays and practice trials. The practice trials had the same sequence of events as the experimental trials, but the practice trials also included a feedback display appearing after the participant responded. The practice session started with 17 practice trials that were easier than the experimental trials: the stimuli contrast was high with no additional noise, and presentation duration and SOAs were longer. The practice session ended with nine practice trials that were identical to the experimental trials (i.e., with the same difficulty level).

An experimental trial started with the "ready" display. Two hundred milliseconds after the participants pressed the spacebar, the sequence of three orientation items was presented at the center of the screen. Each item appeared for 50 ms and was followed by an interstimulus interval of either 150 or 350 ms (i.e., SOA of 200 or 400 ms, respectively). The SOA was the same within a trial but varied randomly and evenly interleaved across trials. The probe circle appeared 500 ms after the last blank display and the white response dot appeared with the first mouse click on the circle. Participants were asked to report only the orientation of the target (i.e., the second item). In the uncrowded condition, only one item appeared (the target) which was followed by the response probe, and the participants had to report the target's orientation. The time course of the uncrowded trials was the same as the crowded trials with 200 ms SOA. Participants adjusted the response dot with their mouse to match the target's orientation and validated their final response by pressing the spacebar. There was no time limit for responding and only accuracy was emphasized. The participants performed 80 trials in each of the three conditions (uncrowded, crowded—200 ms, and crowded—400 ms). The trial type varied randomly across the experimental session. Due to a coding error, the participants completed 241 trials.

Experiment 3

The procedure was identical to Experiment 1 except for the following: Participants sat in a dim room at the lab, approximately 56 cm from the screen, and used a chin rest. There was no "virtual chin-rest" procedure. Once the participants completed the instructions and practice sessions, they performed 100 experimental trials in each of the three conditions, randomly interleaved across the experiment for a total of 300 experimental trials.

Experiment 4

The procedure was identical to Experiment 1 except for the following: There were three conditions: crowded, uncrowded, and control. The crowded and uncrowded conditions were identical to the crowded—200 ms and uncrowded conditions of Experiment 1, respectively. Importantly, the control condition was the same as the crowded—200 ms condition of Experiment 1; however, the first and the last items were a circle without the radius line that indicates the orientation. The participants completed a total of 240 trials.

Mixture-Model Analysis

We analyzed the data using the two-misreport mixture model that includes four parameters (Shechter & Yashar, 2021): (a) The width (*SD*) of a von Mises distribution (normal-circular distribution) of errors that is centered around the target's orientation (i.e., zero error). This parameter reflects the error variance of trials in which the target was encoded. It conveys the encoding precision or the quality of the target representation; the larger the *SD*, the lower the quality of the target representation. (b) The height (g) of a uniform distribution of errors that are due to sheer guessing. This parameter indicates the guessing rate—the frequency of trials in which the target was not registered at all. This is mainly determined by the SNR; the higher the g, the lower the SNR (Agaoglu et al., 2015). (c) The rate of mistakenly reporting the orientation of the first distractor instead of the target (β_1); and (d) the rate of mistakenly reporting the orientation of the last distractor instead of the target (β_2). The β_1 and β_2 parameters are often referred to as substitution errors, and they are modeled by additional von Mises distributions centered on the distractors' orientation.

On each trial, we calculated the error magnitude, that is, the difference between the reported orientation and the target orientation. The error magnitude ranged from −180° to +180°. For the crowded conditions, we used *Memtoolbox* (Suchow et al., 2013) to compare three models:

1. The two-misreport model (Shechter & Yashar, 2021) with its four parameters (detailed above): (a) *SD* which estimates the precision of the target representation; (b) g which estimates the guessing rate; (c) β_1 which estimates the rate of substitution errors involving the first distractor; and (d) β_2 which estimates the rate of substitution errors involving the last distractor:

$$p(\theta) = (1 - g - \beta_1 - \beta_2)\varphi_\sigma(\theta) + g/2\pi + \beta_1\varphi_\sigma(\theta_1^*) + \beta_2\varphi_\sigma(\theta_2^*), \quad (2)$$

   where θ is the response error, and φ denotes the circular analog of the Gaussian distribution (the von Mises distribution) with mean equal to zero (zero error) and standard deviation (*SD*). θ_1* is the error relative to the orientation of the preceding distractor, and θ_2* is the error relative to the orientation of the succeeding distractor.
2. The swap model (Bays et al., 2009) aggregates the contribution of different distractors and, therefore, has only three free parameters—*SD*, g, and β:

$$p(\theta) = (1 - g - \beta)\varphi_\sigma(\theta) + g/2\pi + \beta/m\, \varphi_\sigma(\theta_i^*), \quad (3)$$

   where β is the rate of mistakenly reporting the orientation of one of the distractors, m is the number of distractors (two in our study), and θ_i* is the error relative to the orientation of the ith distractor.
3. The standard model (Zhang & Luck, 2008) does not allow for substitution errors of any kind and therefore it has only two parameters—*SD* and g:

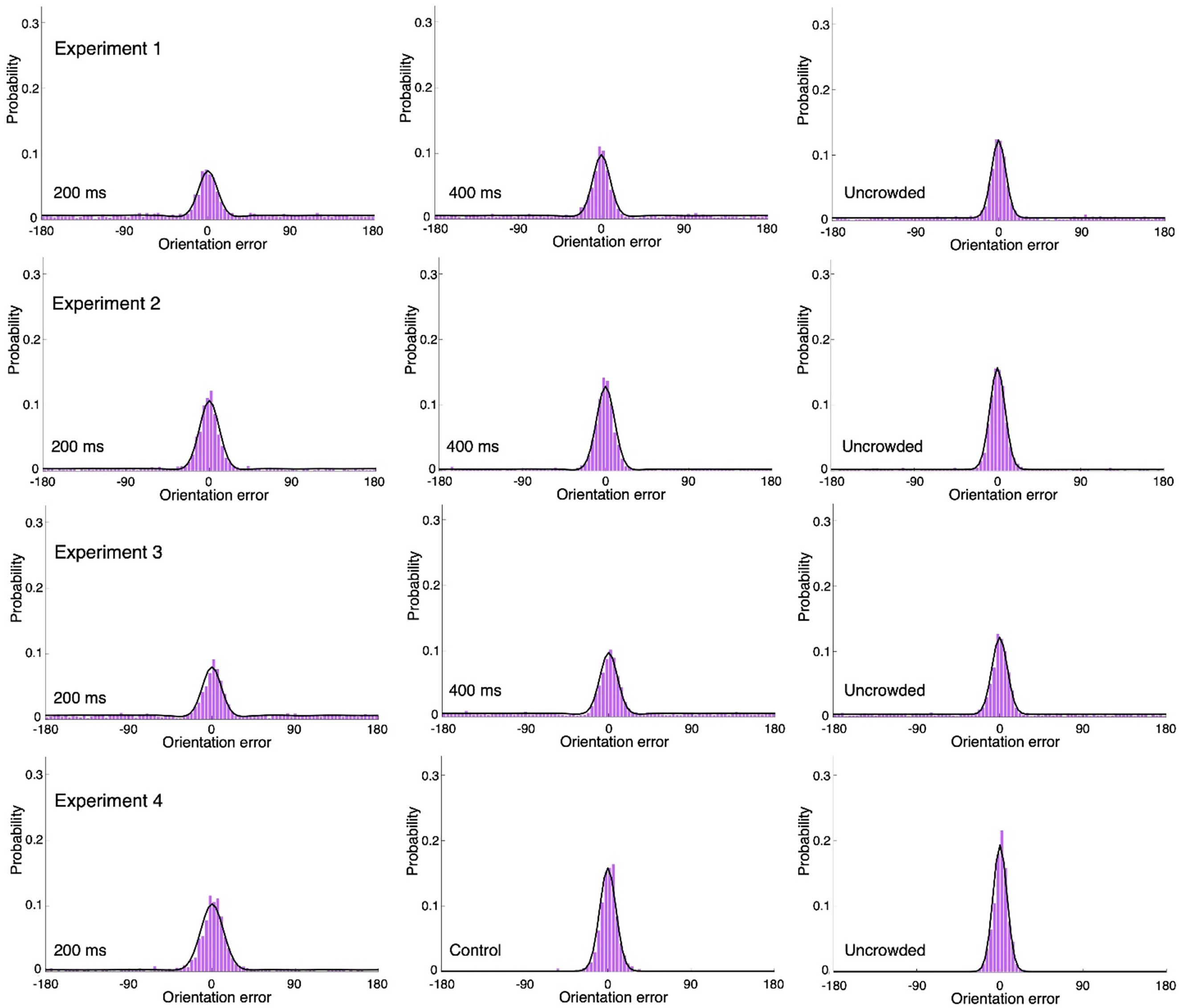
$$p(\theta) = (1 - g)\varphi_\sigma(\theta) + g/2\pi. \quad (4)$$

The models were compared using the Akaike information criterion with correction (AICc) and Bayesian information criterion (BIC), both include a penalty term for each additional model parameter. The two-misreport model (Model 1) received the lowest BIC and AICc scores. The two-misreport model also received lower scores than a version of the swap model that includes a bias term (Suchow et al., 2013). Therefore, we continued the analyses of the crowded conditions of all experiments with this model. Indeed, as can be seen in Figure 2, this model fits the data well. For the uncrowded conditions of Experiments 1–4 and the control condition of Experiment 4, there were no orientation distractors to misreport, and therefore we used the standard model (Model 3) to analyze the data of these conditions.

Transparency and Openness

All the measurements employed in this study, as well as all data exclusions, are reported, and we describe above how we determined our sample size. An interested reader can experience the task at https://run.pavlovia.org/TS.onlineLabExp/temporal-crowding-r2/html/. Materials used in this study and the data are available at the OSF repository (https://osf.io/2eacq). Code and statistical analyses are also available at the OSF repository (https://osf.io/2eacq). None of the analyses or hypotheses were formally preregistered.

Figure 2
Mean Error Distributions and Model Fits (in Black) for the Various Conditions of Experiments 1–4



Note. See the online article for the color version of this figure.

Results

When Tkacz-Domb and Yeshurun (2021) tested the effects of SOA on the four parameters of the model, at the periphery, a significant increase with decreasing SOA was found for the *SD*, β_1, and β_2 parameters, but not for the *g* parameter. The main analyses of the current study examined whether or not a similar pattern will be found at the fovea.

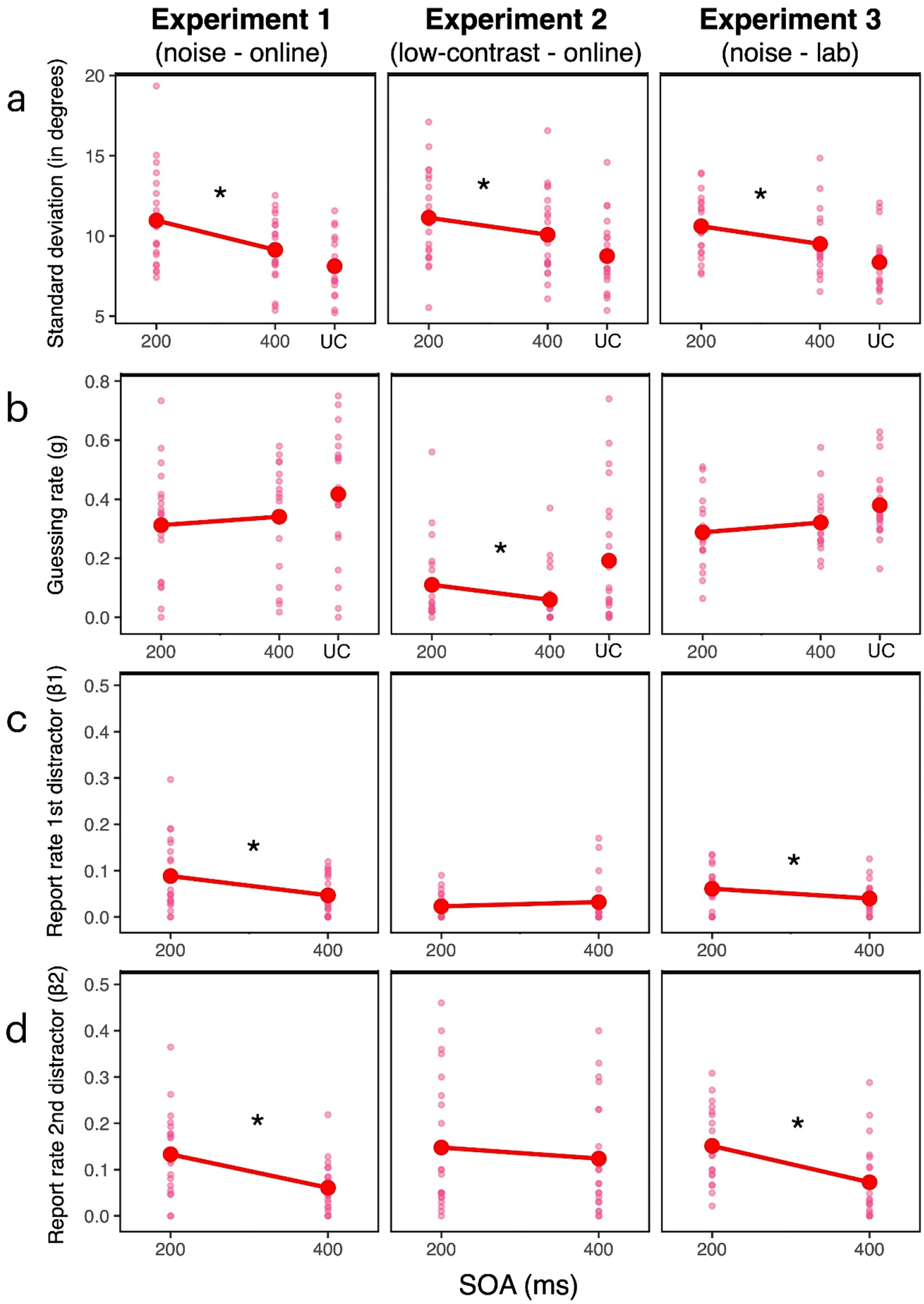
Analyses of Foveal Temporal Crowding Effects: Experiments 1–3

The model parameters were estimated separately for each condition and each participant. The analyses of these parameters were performed using the following R packages: rstatix (Kassambara, 2021), car (Fox & Sanford Weisberg, 2019), and tidyverse (Wickham et al., 2019). We conducted an omnibus two-way mixed-design ANOVA on each parameter with SOA (200 and 400 ms) as the within-subject variable and experiment (1—*online with noise*, 2—*online low contrast*, 3—*lab with noise*) as the between-subject variable (Figure 3).

Precision of Target Representation (SD)

Starting with the *SD* parameter, only a main effect of SOA emerged, $F(1, 61) = 21.44$, $p < .0001$, $\eta_p^2 = .26$, the *SD* was significantly larger when the SOA was shorter, suggesting that temporal crowding decreases the precision of the target representation. Subsequent one-way (SOA) repeated-measures ANOVA conducted for each experiment confirmed that the effect of SOA on the *SD* was significant with all three experiments, Experiment 1, $F(1, 20) =$

Figure 3
The Averages (Large Red Circles) of the Estimated Parameters for Experiments 1–3



Note. (a) *SD*, (b) *g*, (c) β_1, and (d) β_2 as a function of SOA in the crowded condition and the UC condition of each experiment (1–3). The smaller (pink) data points depict individual data. Asterisk "*" indicates a significant ($p < .05$) difference between the SOAs. SOA = stimulus onset asynchrony; UC = uncrowded condition. See the online article for the color version of this figure.

11.94, $p = .002$, $\eta_p^2 = .37$, Experiment 2, $F(1, 22) = 5.44$, $p = .029$, $\eta_p^2 = .20$, Experiment 3, $F(1, 19) = 4.65$, $p = .044$, $\eta_p^2 = .20$. This finding is consistent with the characteristics of temporal crowding at the periphery and further demonstrates its robustness—a similar degradation of the quality of the target's representation was found regardless of whether the experiment was conducted online or in the lab and regardless of the method used to adjust task difficulty for central presentation. Also consistent with temporal crowding at the periphery is the fact that in all our experiments the *SD* for the longer SOA of 400 ms was significantly larger than the *SD* in the uncrowded condition, Experiment 1, $t(20) = 2.24$, $p = .037$, Cohen's $d = .49$, Experiment 2, $t(22) = 3.57$, $p = .002$, Cohen's $d = .74$, Experiment 3, $t(19) = 3.84$, $p = .001$, Cohen's $d = .86$, confirming that the long-lasting degradation of the target representation, brought about by temporal crowding, is also present at the fovea. The main effect of experiment and the SOA × experiment interaction were not significant, $F(2, 61) = .49$, $p = .614$, $\eta_p^2 = .02$, $F(2, 61) = .76$, $p = .472$, $\eta_p^2 = .02$, respectively.

Substitution Errors (β_1, β_2)

A significant main effect of SOA also emerged for the two parameters—β_1 and β_2—that reflect substitution errors, β_1, $F(1, 61) = 8.24$, $p = .006$, $\eta_p^2 = .12$; β_2, $F(1, 61) = 43.01$, $p < .0001$, $\eta_p^2 = .41$, but with both this main effect was qualified by an SOA × experiment interaction, β_1, $F(2, 61) = 5.83$, $p = .005$, $\eta_p^2 = .16$; β_2, $F(2, 61) = 3.91$, $p = .025$, $\eta_p^2 = .11$. The nature of this interaction was revealed by one-way (SOA) repeated-measures ANOVA conducted for each experiment: For both parameters the SOA effect was significant with Experiments 1 and 3 but not Experiment 2, β_1: Experiment 1, $F(1, 20) = 7.11$, $p = .015$, $\eta_p^2 = .26$; Experiment 2, $F(1, 22) = 1.76$, $p = .198$, $\eta_p^2 = .07$; Experiment 3: $F(1, 19) = 6.98$, $p = .016$, $\eta_p^2 = .27$; β_2: Experiment 1, $F(1, 20) = 21.07$, $p = .0002$, $\eta_p^2 = .51$; Experiment 2, $F(1, 22) = 2.25$, $p = .148$, $\eta_p^2 = .09$; Experiment 3, $F(1, 19) = 31.67$, $p < .0001$, $\eta_p^2 = .63$. Thus, when task difficulty was adjusted by embedding the stimuli in noise, the effect of SOA at the center replicated the SOA effect found at the periphery: Substitution errors were significantly more frequent for shorter than longer SOA, suggesting that temporal crowding increases source confusion with both the preceding and succeeding distractors. However, when the stimuli contrast was reduced substitution errors were similar regardless of the SOA. With the succeeding distractor (β_2), the rate of substitution errors was high for both SOAs (Figure 3d). This may be due to the difficulty to detect the low-contrast stimuli. Perhaps on some of the trials, regardless of their SOA, some items in the sequence were not detected due to their low contrast. This made it difficult to determine which item is the target, resulting in many substitution errors even when the interitem interval was long. Regarding the preceding distractor (β_1), the lack of SOA effect may simply reflect a floor effect (Figure 3c). Indeed, the omnibus ANOVA for the first distractor (β_1) also revealed a main effect of experiment, $F(2, 61) = 5.17$, $p = .008$, $\eta_p^2 = .15$, the rate of substitution errors in Experiment 2 was significantly lower than that of Experiment 1, $t(42) = 3.01$, $p = .004$, Cohen's $d = .91$ or Experiment 3, $t(41) = 2.16$, $p = .036$, Cohen's $d = .66$. There was no significant difference between Experiments 1 and 3, $t(39) = 1.21$, $p = .233$, Cohen's $d = .38$. For β_2 the main effect of the experiment was not significant, $F(2, 61) = .98$, $p = .382$, $\eta_p^2 = .03$.

Guessing Rate (g)

The omnibus ANOVA for the *g* parameter revealed a main effect of the experiment, $F(2, 61) = 24.57$, $p < .0001$, $\eta_p^2 = .45$. The guessing rate was significantly lower in Experiment 2 than in Experiment 1: $t(42) = 5.71$, $p < .0001$, Cohen's $d = 1.72$ and Experiment 3: $t(41) = 7.16$, $p < .0001$, Cohen's $d = 2.19$. There was no significant difference between Experiments 1 and 3, $t(39) = .52$, $p = .606$, Cohen's $d = .16$. As mentioned above, the guessing rate reflects the SNR. Thus, the higher guessing rate in Experiments 1 and 3 in comparison to Experiment 2 may suggest that the SNR was lower when the stimuli were embedded in external noise than when their contrast was low. However, it is important to note that the source of noise that led to a given SNR, in these two cases, is not quite comparable. In Experiments 1 and 3, the main source of noise was the external noise added to each stimulus in the sequence. The contrast between the overall gray background and each stimulus (orientation item + noise) was relatively high and the contribution of internal noise was likely negligible. Accordingly, the mere detection of each item of the sequence was not particularly hard in these experiments. Rather it was hard to discern the orientation item from its corresponding external noise patch. In Experiment 2, the main source of noise was internal noise because no external noise was added and the contrast between each orientation item and the overall gray background was low. Accordingly, it is likely that in some of the trials of Experiment 2 not all three items were detected. This makes the experimental setting of Experiments 1 and 3 more similar to that employed with peripheral presentation (Tkacz-Domb & Yeshurun, 2021) because the items had relatively high contrast and it was not particularly hard to detect them.

The omnibus ANOVA for the *g* parameter also revealed a significant SOA × experiment interaction, $F(2, 61) = 4.54$, $p = .015$, $\eta_p^2 = .13$. Subsequent one-way repeated-measures ANOVA conducted for each experiment suggested that the interaction is due to the emergence of a significant SOA effect with Experiment 2—the guessing rate was higher with the shorter SOA, $F(1, 22) = 14.52$, $p = .001$, $\eta_p^2 = .40$, but not with Experiments 1 and 3, Experiment 1, $F(1, 20) = 1.15$, $p = .296$, $\eta_p^2 = .05$; Experiment 3, $F(1, 19) = 1.53$, $p = .232$, $\eta_p^2 = .07$. The lack of an SOA effect on the guessing rate observed in Experiments 1 and 3 echoes the findings of Tkacz-Domb and Yeshurun (2021). None of the three experiments conducted in that study revealed an effect of temporal crowding on the guessing rate. Thus, the significant SOA effect that emerged in Experiment 2 of the current study is the exception, and it is probably related to the items' low contrast.

Comparison of Foveal and Peripheral Temporal Crowding

The analyses above demonstrate that temporal crowding occurs with central presentation. In this section, we examine whether the magnitude of temporal crowding is decreased with central presentation, as is the case with spatial crowding. To that end, we compare the effects of SOA on each of the parameters with central and peripheral presentation. For the central presentation, we chose Experiment 1 from the current study because it is similar to Experiment 3, and unlike Experiment 2, its high-contrast stimuli match those of Tkacz-Domb and Yeshurun (2021). For the peripheral presentation, we chose Experiment 2 from Tkacz-Domb and Yeshurun (2021) because that experiment included the SOAs of 220 and 420 ms which are closest to the SOAs employed in the current study. Only these two SOAs from the crowded condition of Experiment 2 (Tkacz-Domb & Yeshurun, 2021) were included in the comparison, and for the sake of simplicity, hereafter we ignore the 20 ms difference and refer to the SOAs of both experiments as 200 and 400 ms. We note that stimuli duration was 50 ms in the "central" experiment and 20 ms in the "peripheral" experiment, but Tkacz-Domb and Yeshurun (2021) showed that the pattern of the SOA effects on the different parameters is similar for stimuli duration of 20 and 75 ms. Additionally, with the peripheral experiment, there was spatial uncertainty before the onset of the first item because the items could appear to the right or the left of fixation (though, following the first item there was no longer uncertainty because all the items were presented to the same location). With the central presentation, there was no such spatial uncertainty.

We conducted a two-way mixed-design ANOVA on each parameter with SOA (200 and 400 ms) as the within-subject variable and position within the visual field (center vs. periphery) as the between-subject variable (Figure 4). Not surprisingly, a significant main effect of SOA was found with the *SD*, β_1, and β_2 parameters, *SD*, $F(1, 36) = 13.77$, $p = .001$, $\eta_p^2 = .28$; β_1, $F(1, 36) = 19.10$, $p = .0001$, $\eta_p^2 = .35$; β_2, $F(1, 36) = 35.82$, $p < .0001$, $\eta_p^2 = .50$, but not with the *g* parameter, $F(1, 36) = 1.40$, $p = .244$, $\eta_p^2 = .04$. A significant main effect of position within the visual field emerged for the *SD*, β_1, and *g* parameters. The *SD* was larger at the periphery than the center, $F(1, 36) = 87.61$, $p < .0001$, $\eta_p^2 = .71$, as expected given the lower spatial resolution at the periphery. The β_1 was also larger at the periphery than at the center, $F(1, 36) = 24.59$, $p < .0001$, $\eta_p^2 = .41$. This finding may be related to the fact that with the peripheral experiment, there was uncertainty regarding the position of the first distractor, and this led to more confusion between this item and the target. In contrast, the *g* parameter was larger in the central than peripheral presentation, $F(1, 36) = 4.73$, $p = .036$, $\eta_p^2 = .12$, likely because the central items were embedded in noise (i.e., lower SNR results in a higher guess rate). There was no significant effect with the β_2 parameter, $F(1, 36) = 1.07$, $p = .307$, $\eta_p^2 = .03$.

Critically, with the *SD*, β_2, and *g* parameters the interaction between SOA and position within the visual field was not significant, *SD*, $F(1, 36) = .22$, $p = .642$, $\eta_p^2 = .01$; β_2, $F(1, 36) = .98$, $p = .328$, $\eta_p^2 = .03$; *g*, $F(1, 36) = .04$, $p = .836$, $\eta_p^2 = .001$. Only with the β_1 parameter the interaction was significant, $F(1, 36) = 4.52$, $p = .040$, $\eta_p^2 = .11$, however, as can be seen in Figure 4c, this interaction may simply be due to a floor effect because with central presentation the overall rate of mistakenly reporting the first distractor was very low. This suggests that regardless of the stimuli position (central or peripheral), temporal crowding impaired the encoding precision and increased substitution errors but did not increase the guessing rate. Moreover, for the most part, the magnitude of temporal crowding is similar at the center and the periphery of the visual field.

Experiment 4

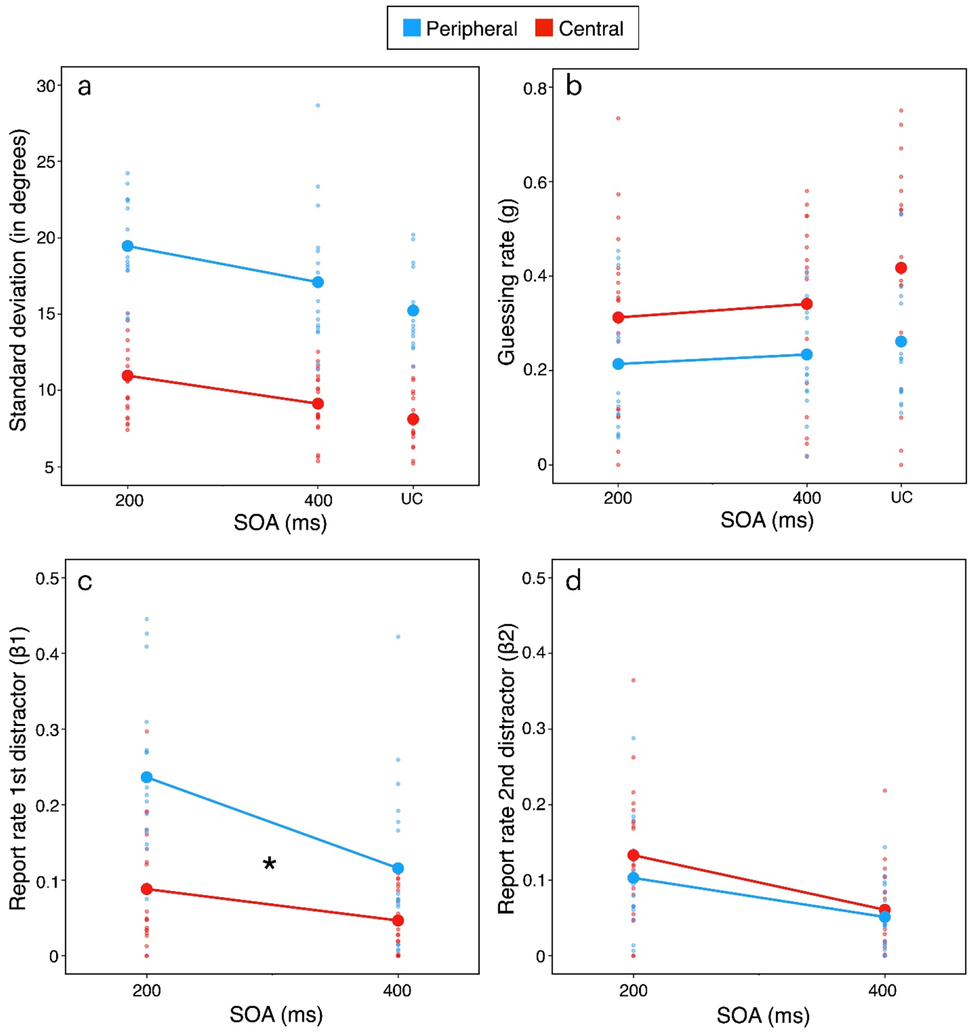
In Experiments 1–3, the crowded and uncrowded trials differed in two aspects. First, an uncrowded trial included a single event, while a crowded trial included three events, each involving an onset and offset of a stimulus. Second, in the uncrowded trials, only the target included features of the relevant dimension (i.e., orientation), while in the crowded trials, the distractors also included orientation information which potentially introduced response competition. Experiment 4 examined the relative contribution of each of these aspects to the effect of temporal crowding on the quality of the target representation (*SD*). The experiment included three conditions. Two conditions were similar to those of Experiment 1: a crowded condition with an SOA of 200 ms and an uncrowded condition. In the third control condition, the distractors were similar circles embedded in noise, only without the orientation line (Figure 1d). Thus, a trial in this condition included three events, but only the target contained orientation information. A one-way (condition: uncrowded, control, crowded) repeated-measures ANOVA performed on the *SD* parameter revealed a significant condition effect, $F(2, 44) = 31.58$, $p < .0001$, $\eta_p^2 = .589$ (Figure 5). The *SD* in the control condition was significantly larger than the *SD* in the uncrowded condition, $t(22) = 3.28$, $p = .003$, Cohen's $d = .68$. This finding indicates that the distractors do not need to share the task-relevant dimension with the target in order to generate temporal interference. Interestingly, the *SD* in the crowded condition was also significantly larger than the *SD* in the control condition, $t(22) = 6.62$, $p < .0001$, Cohen's $d = 1.38$. This finding suggests that when the distractors contain features in the task-relevant dimension they further interfere with the encoding of the target. To test which of these components of interference is more dominant in the overall effect of temporal crowding, we calculated for each participant, the difference between the *SD*s of the uncrowded and control conditions (reflecting interference generated by the mere presence of distractors) and the difference between the *SD*s of the crowded and control conditions (reflecting interference generated by the presence of task-relevant features in the distractors). A paired *t* test comparing these two aspects of interference indicated that the crowded-control difference ($M = 4.97$, $SD = 3.60$) was significantly larger, $t(22) = 7.2073$, $p < .0001$, Cohen's $d = 1.50$, than the uncrowded-control difference ($M = 1.69$, $SD = 2.47$). Thus, although both aspects of temporal crowding contribute to the long-lasting degradation of the target representation, the presence of distractors that share the task-relevant dimension with the target plays a more prominent role.

General Discussion

This study examined whether temporal crowding also occurs at the center of the visual field, and if so whether its magnitude is reduced. To that end, a sequence of three stimuli, separated by relatively long SOAs, was presented to the center of the screen. To adjust task difficulty for foveal presentation the stimuli were embedded in noise (Experiments 1, 3, and 4) or their contrast was reduced (Experiment 2). The SOA varied between trials, and the task was an orientation estimation task. The error distributions were analyzed using the two-misreport mixture model (e.g., Shechter & Yashar, 2021). These analyses revealed clear evidence of temporal crowding with central vision. Starting with the precision of target representation, temporal crowding decreased this precision in all three experiments, and the magnitude of this impairment was similar to peripheral temporal crowding. Combining this pattern of results and that found by Tkacz-Domb and Yeshurun (2021) shows that reduced precision with shorter SOAs was found regardless of position within the visual field (central vs. peripheral), stimuli duration (20 vs. 75 ms), target-distractors similarity (similar vs. dissimilar),

Figure 4

The Averages (Large Circles) of the Estimated Parameters for Central and Peripheral Presentation



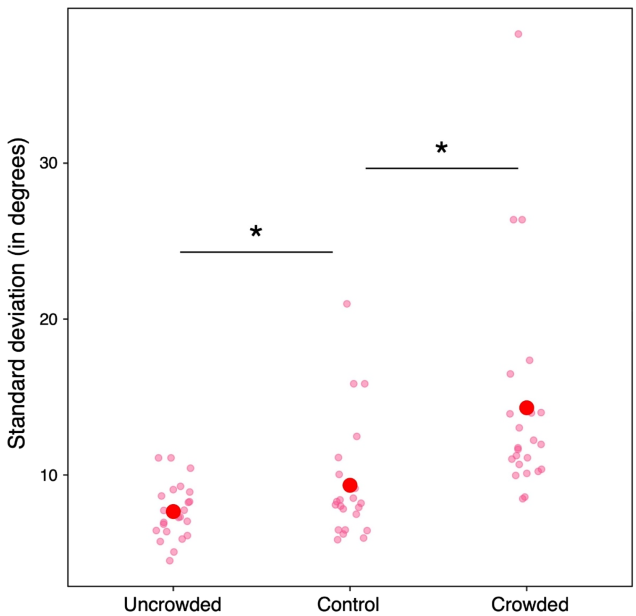
Note. (a) *SD*, (b) *g*, (c) β_1, and (d) β_2 as a function of SOA in the crowded and UC conditions when the sequence of stimuli was presented at the center (in red/dark gray) and when the sequence of stimuli was presented at the periphery (in blue/light gray). The smaller data points depict individual data. Asterisk "*" indicates a significant ($p < .05$) interaction. The central presentation data is taken from the current Experiment 1 and the peripheral presentation data is adapted from "Temporal Crowding Is a Unique Phenomenon Reflecting Impaired Target Encoding Over Large Temporal Intervals," by S. Tkacz-Domb and Y. Yeshurun, 2021, *Psychonomic Bulletin & Review*, *28*(6), pp. 1885–1893 (https://doi.org/10.3758/s13423-021-01943-8). Copyright 2021 by the American Psychological Association. UC = uncrowded condition; SOA = stimulus onset asynchrony. See the online article for the color version of this figure.

the method employed to increase task difficulty (added external noise vs. low contrast), whether or not there was spatial uncertainty before the onset of the sequence, and the platform for data collection (online vs. lab). Thus, it becomes clear that the degradation of target representation (i.e., reduced precision) is the most noticeable and robust characteristic of temporal crowding.

The impaired precision of target representation is not only a prominent characteristic of temporal crowding but also a unique characteristic of temporal crowding that distinguishes between it and other types of interference. For instance, Ester et al. (2014) used similar stimuli and procedure to examine spatial crowding but did not find effects of target-distractor distance (the spatial analog of SOA) on precision. Similarly, Agaoglu et al. (2015) employed an orientation estimation task and mixture-model analysis to examine various types of ordinary masking, including pattern masking by structure, which is the type of masking most relevant for comparison with our stimuli because it does not spatially overlap the target and shares some of its features. Unlike our findings, they mainly found an effect of SOA on the guessing rate but not on the encoding precision. Moreover, the fact that the magnitude of the detrimental

Figure 5

The Averages of the Estimated SD *Parameter for Each of the Conditions of Experiment 4: Uncrowded—A Single Item: the Target; Control—Three Items: the Target and Two Circles (SOA = 200 ms); Crowded—Three Items: the Target and Two Circles With an Oriented Line (SOA = 200 ms)*



Note. The small pink data points depict individual data. Asterisk "*" indicates a significant ($p < .05$) difference. SOA = stimulus onset asynchrony. See the online article for the color version of this figure.

effect of temporal crowding on the target's precision was similar with central and peripheral presentations is also in striking difference from spatial crowding and forward/backward masking whose detrimental effect on performance is considerably smaller in central regions (e.g., Breitmeyer & Ganz, 1976; Coates et al., 2018; Francis, 2003; Matthews, 1973; Strasburger, 2020). Therefore, the current study portrays impaired encoding precision as the hallmark of temporal crowding and provides further support to the conclusion that temporal crowding is qualitatively different from other known types of interitem impairment.

The finding that temporal crowding does not vary with respect to retinotopic position is of particular theoretical importance because it helps us better understand the mechanisms of temporal crowding. It suggests that the degradation of target representation by temporal crowding likely reflects processes taking place at higher visual areas for which differences between different retinal positions like acuity, spatial/temporal summation, and receptive/integration field size, no longer play a major role. Additional insights regarding the mechanisms underlying temporal crowding are provided by Experiment 4. This experiment examined two possible sources of interference: the mere presence of crowding stimuli (i.e., the distractors) and the fact that the distractors also include orientation information. We found that both sources of interference contribute to the detrimental effect of temporal crowding on the precision of target encoding; however, the latter source of interference has a more substantial contribution. The finding that the mere presence of distractors is enough to degrade precision, even when the distractors do not include task-relevant features, suggests that response competition is not the sole mechanism of temporal crowding. The finding that the precision impairment was larger when both the distractors and the target included orientation information suggests that a significant part of temporal crowding reflects long-lasting interference between neuronal populations that specifically encode orientation information.

Combining the findings that the magnitude of temporal crowding is not reduced in the fovea or affected by the presence/absence of spatial uncertainty about the sequence location, together with the finding that the detrimental effect on encoding precision persists beyond an SOA of 400 ms even at the fovea, further qualifies the widespread notion that a representation of visual information is generated very fast. This common notion is based on various findings showing that about 100–150 ms following the onset of a centrally presented visual stimulus the participants can say quite a lot about it, like indicating the basic category of a natural scene, determining whether or not it contained an animal, and reporting several large objects incorporated in the scene (e.g., Bacon-Macé et al., 2005; Castelhano & Henderson, 2008; Fei-Fei et al., 2007; Greene & Oliva, 2009; Grill-Spector & Kanwisher, 2005; Thorpe et al., 1996). Our task-relevant stimulus was considerably simpler than an image of a natural scene, yet our results and those of Tkacz-Domb and Yeshurun (2021) indicated that the presence of task-irrelevant items degraded the quality of the target representation even when these items appeared more than 400 ms before/after the target. Thus, an initial volatile visual representation may be generated fast, but our current findings suggest that even with central vision, a robust and stable representation requires considerably longer processing time than the common belief.

Our findings of long-lasting impairment of the target representation are consistent with other studies that have shown temporal interactions over long intervals (Otto et al., 2009; Scharnowski et al., 2009), but they are inherently different from the attentional blink phenomenon (Raymond et al., 1992). Attentional blink studies usually present a central fast stream of high-contrast stimuli (SOAs around 100 ms), and the participants have to report two targets. The term "attentional blink" refers to the impaired identification of the second target, and this impairment is typically attributed to the need to encode the first target (e.g., Chun & Potter, 1995; Raymond et al., 1992; for a recent review, see Snir & Yeshurun, 2017). In contrast, with the temporal crowding task, the participants are required to report only a single target, and therefore the observed impairment cannot be attributed to actively encoding another target or any of the other accounts of the attentional blink as they all require the presence of an additional target. Interestingly, attentional blink studies do not usually find impairment for the first target, even though other task-irrelevant items precede and succeed it. This is likely due to the central presentation of high-contrast stimuli. Anyhow, the range of SOAs employed in those studies is within the range of ordinary masking, not temporal crowding, and as detailed above, these are two different phenomena.

Another unlikely mediator of temporal crowding is memory capacity limitation. The stimuli were within capacity limits and the participants were required to remember only a single item—the target. Critically, they knew in advance that there was no need to memorize the entire sequence, as the target was always the second item of the sequence. This conclusion is further supported by the fact that temporal crowding was not eliminated even when only a single oriented line was presented (Experiment 4).

Lastly, when considering guessing rate and substitution errors, the results are less robust than those of encoding precision and depend on the method used to increase task difficulty. When the items were embedded in noise (Experiments 1 and 3), the results were very similar to those observed with peripheral presentation. Temporal crowding did not affect the guessing rate, and with both distractors, the rate of substitution errors varied as a function of SOA. Additionally, when comparing the magnitude of temporal crowding on substitution rate at the periphery and the center, a different pattern emerged for the first and last distractors. For the last distractor, the magnitude of temporal crowding was similar regardless of position within the visual field, but for the first distractor, temporal crowding was more pronounced under peripheral presentation. As mentioned above, these differences might be due to the overall low rate of mistakenly reporting the first distractor, which may be related to differences in spatial uncertainty between the two tasks regarding this distractor. Interestingly, when the items' contrast against the gray background was reduced (Experiment 2), the pattern of results differed from that of peripheral presentation—there was no effect of SOA on substitution errors yet an SOA effect on the guessing rate emerged. These differences might reflect difficulties to detect all three events with the low-contrast paradigm. Thus, maybe once detection processes are involved, temporal crowding also affects the SNR.

To summarize, temporal crowding can be found with a central presentation, and the magnitude of its detrimental effect on target encoding is similar to central and peripheral presentations. Moreover, degradation of encoding precision emerged regardless of the experimental platform (online/lab) or method for increasing task difficulty, marking it as the most robust and unique aspect of temporal crowding. Finally, the target encoding was impaired even when the distractors did not include task-relevant features, but the impairment was considerably larger when all items included orientation information. Taken together, these findings show, unequivocally, that even a simple central visual representation is susceptible to interference for much longer timescales than previously thought portraying our visual system as surprisingly volatile.

Constraints on Generality

This study used heterogenic neurotypical samples (aged 18–35) that included a similar number of females and males recruited from the general public through an online platform in Experiments 1, 2, and 4 (English as a first language) or were university students in Experiment 3 (Hebrew as a first language). A similar pattern of results was obtained with both the online platform and the lab setting. Thus, when considering neurotypical adults, the current results are not constrained to a specific target population.

References

Agaoglu, S., Agaoglu, M. N., Breitmeyer, B., & Ogmen, H. (2015). A statistical perspective to visual masking. *Vision Research*, *115*(Pt. A), 23–39. https://doi.org/10.1016/j.visres.2015.07.003

Bacon-Macé, N., Mace, M. J. M., Fabre-Thorpe, M., & Thorpe, S. J. (2005). The time course of visual processing: Backward masking and natural scene categorization. *Vision Research*, *45*(11), 1459–1469. https://doi.org/10.1016/j.visres.2005.01.004

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(2), Article 7. https://doi.org/10.1167/9.10.7

Bonneh, Y. S., Sagi, D., & Polat, U. (2007). Spatial and temporal crowding in amblyopia. *Vision Research*, *47*(14), 1950–1962. https://doi.org/10.1016/j.visres.2007.02.015

Breitmeyer, B. G. (1984). *Visual masking: An integrative approach*. Oxford University Press.

Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review*, *83*(1), 1–36. https://doi.org/10.1037/0033-295X.83.1.1

Breitmeyer, B. G., & Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics*, *62*(8), 1572–1595. https://doi.org/10.3758/BF03212157

Breitmeyer, B. G., & Ogmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision*. Oxford University Press.

Castelhano, M. S., & Henderson, J. M. (2008). The influence of color on the perception of scene gist. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(3), 660–675. https://doi.org/10.1037/0096-1523.34.3.660

Champely, S. (2020). *Basic functions for power analysis* (R package Version 1.3-0) [Computer software]. https://cran.r-project.org/package=pwr

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 109–127. https://doi.org/10.1037/0096-1523.21.1.109

Coates, D. R., Levi, D. M., Touch, P., & Sabesan, R. (2018). Foveal crowding resolved. *Scientific Reports*, *8*(1), Article 9177. https://doi.org/10.1038/s41598-018-27480-4

Enns, J. T. (2004). Object substitution and its relation to other forms of visual masking. *Vision Research*, *44*(12), 1321–1331. https://doi.org/10.1016/j.visres.2003.10.024

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*(9), 345–352. https://doi.org/10.1016/S1364-6613(00)01520-5

Ester, E. F., Klee, D., & Awh, E. (2014). Visual crowding cannot be wholly explained by feature pooling. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1022–1033. https://doi.org/10.1037/a0035377

Fei-Fei, L., Iyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real-world scene? *Journal of Vision*, *7*(1), Article 10. https://doi.org/10.1167/7.1.10

Fox, J., & Sanford Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Sage Publications. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Francis, G. (2003). Developing a new quantitative account of backward masking. *Cognitive Psychology*, *46*(2), 198–226. https://doi.org/10.1016/S0010-0285(02)00521-2

Gorea, A. (1987). Masking efficiency as a function of stimulus onset asynchrony for spatial frequency detection and identification. *Spatial Vision*, *2*(1), 51–60. https://doi.org/10.1163/156856887X00051

Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science*, *20*(4), 464–472. https://doi.org/10.1111/j.1467-9280.2009.02316.x

Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science*, *16*(2), 152–160. https://doi.org/10.1111/j.0956-7976.2005.00796.x

Jimenez, M., Kimchi, R., & Yashar, A. (2022). Mixture-modeling approach reveals global and local processes in visual crowding. *Scientific Reports*, *12*(1), Article 6726. https://doi.org/10.1038/s41598-022-10685-z

Kassambara, A. (2021). *rstatix: Pipe-friendly framework for basic statistical tests* (R package Version 0.7.0) [Computer software]. https://CRAN.R-project.org/package=rstatix

Kewan-Khalayly, B., & Yashar, A. (2022). The role of spatial attention in crowding and feature binding. *Journal of Vision*, *22*(13), Article 6. https://doi.org/10.1167/jov.22.13.6

Lev, M., Yehezkel, O., & Polat, U. (2014). Uncovering foveal crowding? *Scientific Reports*, *4*(1), Article 4067. https://doi.org/10.1038/srep04067

Li, Q., Joo, S. J., Yeatman, J. D., & Reinecke, K. (2020). Controlling for participants' viewing distance in large-scale, psychophysical online experiments using a virtual chinrest. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-019-56847-4

Malania, M., Herzog, M. H., & Westheimer, G. (2007). Grouping of contextual elements that affect vernier thresholds. *Journal of Vision*, *7*(2), Article 1. https://doi.org/10.1167/7.2.1

Manassi, M., & Whitney, D. (2018). Multi-level crowding and the paradox of object recognition in clutter. *Current Biology*, *28*(3), R127–R133. https://doi.org/10.1016/j.cub.2017.12.051

Matthews, M. L. (1973). Locus of presentation and the selective masking effect. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *27*(3), 343–349. https://doi.org/10.1037/h0082485

Morys-Carter, W. L. (2021, May 18). *ScreenScale* [Computer software]. Pavlovia. https://doi.org/10.17605/OSF.IO/8FHQK

Otto, T. U., Ögmen, H., & Herzog, M. H. (2009). Feature integration across space, time, and orientation. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1670–1686. https://doi.org/10.1037/a0015798

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pelli, D. G. (2008). Crowding: A cortical constraint on object recognition. *Current Opinion in Neurobiology*, *18*(4), 445–451. https://doi.org/10.1016/j.conb.2008.09.008

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 849–860. https://doi.org/10.1037//0096-1523.18.3.849

Rosenholtz, R. (2016). Capabilities and limitations of peripheral vision. *Annual Review of Vision Science*, *2*(1), 437–457. https://doi.org/10.1146/annurev-vision-082114-035733

Scharnowski, F., Rüter, J., Jolij, J., Hermens, F., Kammer, T., & Herzog, M. H. (2009). Long-lasting modulation of feature integration by transcranial magnetic stimulation. *Journal of Vision*, *9*(6), Article 1. https://doi.org/10.1167/9.6.1

Shechter, A., & Yashar, A. (2021). Mixture model investigation of the inner–outer asymmetry in visual crowding reveals a heavier weight towards the visual periphery. *Scientific Reports*, *11*(1), Article 2116. https://doi.org/10.1038/s41598-021-81533-9

Snir, G., & Yeshurun, Y. (2017). Perceptual episodes, temporal attention, and the role of cognitive control: Lessons from the attentional blink. In C. Howard (Ed.), Progress in brain research: Vol. 236. *Temporal sampling and representation updating* (pp. 53–74). Academic Press.

Strasburger, H. (2020). . Seven myths on crowding and peripheral vision. *i-Perception*, *11*(3), Article 204166952091305. https://doi.org/10.1177/2041669520913052

Suchow, J. W., Brady, T. F., Fougnie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision*, *13*(10), Article 9. https://doi.org/10.1167/13.10.9

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522. https://doi.org/10.1038/381520a0

Tkacz-Domb, S., & Yeshurun, Y. (2017). Spatial attention alleviates temporal crowding, but neither temporal nor spatial uncertainty are necessary for the emergence of temporal crowding. *Journal of Vision*, *17*(3), Article 9. https://doi.org/10.1167/17.3.9

Tkacz-Domb, S., & Yeshurun, Y. (2021). Temporal crowding is a unique phenomenon reflecting impaired target encoding over large temporal intervals. *Psychonomic Bulletin & Review*, *28*(6), 1885–1893. https://doi.org/10.3758/s13423-021-01943-8

Virsu, V., & Rovamo, J. (1979). Visual resolution, contrast sensitivity, and the cortical magnification factor. *Experimental Brain Research*, *37*(3), 475–494. https://doi.org/10.1007/BF00236818

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), Article 1686. https://doi.org/10.21105/joss.01686

Yashar, A., Wu, X., Chen, J., & Carrasco, M. (2019). Crowding and binding: Not all feature dimensions behave in the same way. *Psychological Science*, *30*(10), 1533–1546. https://doi.org/10.1177/0956797619870779

Yeshurun, Y., Rashal, E., & Tkacz-Domb, S. (2015). Temporal crowding and its interplay with spatial crowding. *Journal of Vision*, *15*(3), Article 11. https://doi.org/10.1167/15.3.11

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235. https://doi.org/10.1038/nature06860

Received December 6, 2022
Revision received July 7, 2023
Accepted August 26, 2023 ■