Short Report

Association for Psychological Science

Letters Lost: Capturing Appearance in Crowded Peripheral Vision Reveals a New Kind of Masking

Psychological Science
2019, Vol. 30(7) 1082–1086

DOI: 10.1177/0956797619847166
www.psychologicalscience.org/PS
SAGE

Bilge Sayim[1,2] and Henry Taylor[3]
[1]Institute of Psychology, University of Bern; [2]Sciences Cognitives et Sciences Affectives (SCALab), Unité Mixte de Recherche (UMR) 9193, Centre National de la Recherche Scientifique (CNRS), Université de Lille; and [3]Department of Philosophy, University of Birmingham

Abstract
Peripheral vision is strongly limited by *crowding*, the deleterious influence of flanking items on target perception. Distinguishing what is seen from what is merely inferred in crowding is difficult because task demands and prior knowledge may influence observers' reports. Here, we used a standard identification task in which participants were susceptible to these influences, and to minimize them, we used a free-report-and-drawing paradigm. Three letters were presented in the periphery. In Experiment 1, 10 participants were asked to identify the central target letter. In Experiment 2, 25 participants freely named and drew what they saw. When three identical letters were presented, performance was almost perfect in Experiment 1, but it was very poor in Experiment 2, in which most participants reported only two letters. Our study reveals limitations of standard crowding paradigms and uncovers a hitherto unrecognized effect that we call *redundancy masking*.

Keywords
visual perception, crowding, appearance

Received 5/1/18; Revision accepted 3/5/19

We usually have the mistaken impression of unconstrained, high-resolution access to objects within our entire visual field. However, the largest part of the visual field is peripheral and strongly limited by *crowding*, the deleterious influence of neighboring stimuli on target perception (Bouma, 1970; Levi, 2008). For example, peripheral letter identification deteriorates when the target is surrounded by flanking letters (Fig. 1a). Crowding is generally stronger when the target and the flankers are near each other (Toet & Levi, 1992), similar (Kooi, Toet, Tripathy, & Levi, 1994), and grouped together (Herzog, Sayim, Chicherov, & Manassi, 2015; Sayim, Westheimer, & Herzog, 2010).

In a special case of crowding, *identity crowding* (Block, 2012), the target and the flankers are the same (Fig. 1a). The strength of target disruption in identity crowding is poorly understood. On one hand, the disruptive effects of crowding are stronger when the target and flankers are similar, so we might expect that target identification in identity-crowding conditions is difficult. On the other hand, it was recently proposed that target-identification performance in identity-crowding conditions is superior to normal crowding (Block, 2012; cf. Taylor & Sayim, 2018). To evaluate these two hypotheses, we need an experimental paradigm that can test what is genuinely seen in identity crowding.

The investigation of identity crowding poses unique methodological challenges. Because the target and the flankers are the same, it is difficult to separate target reports from flanker reports, and crucially, reporting a flanker is a "correct" response. Furthermore, observers often have prior stimulus knowledge, for example, because they are informed that three letters are presented. Here, using a standard crowding paradigm, we found

Corresponding Author:
Bilge Sayim, University of Bern, Institute of Psychology, Fabrikstrasse 8, 3012 Bern, Switzerland
E-mail: bilge.sayim@univ-lille.fr

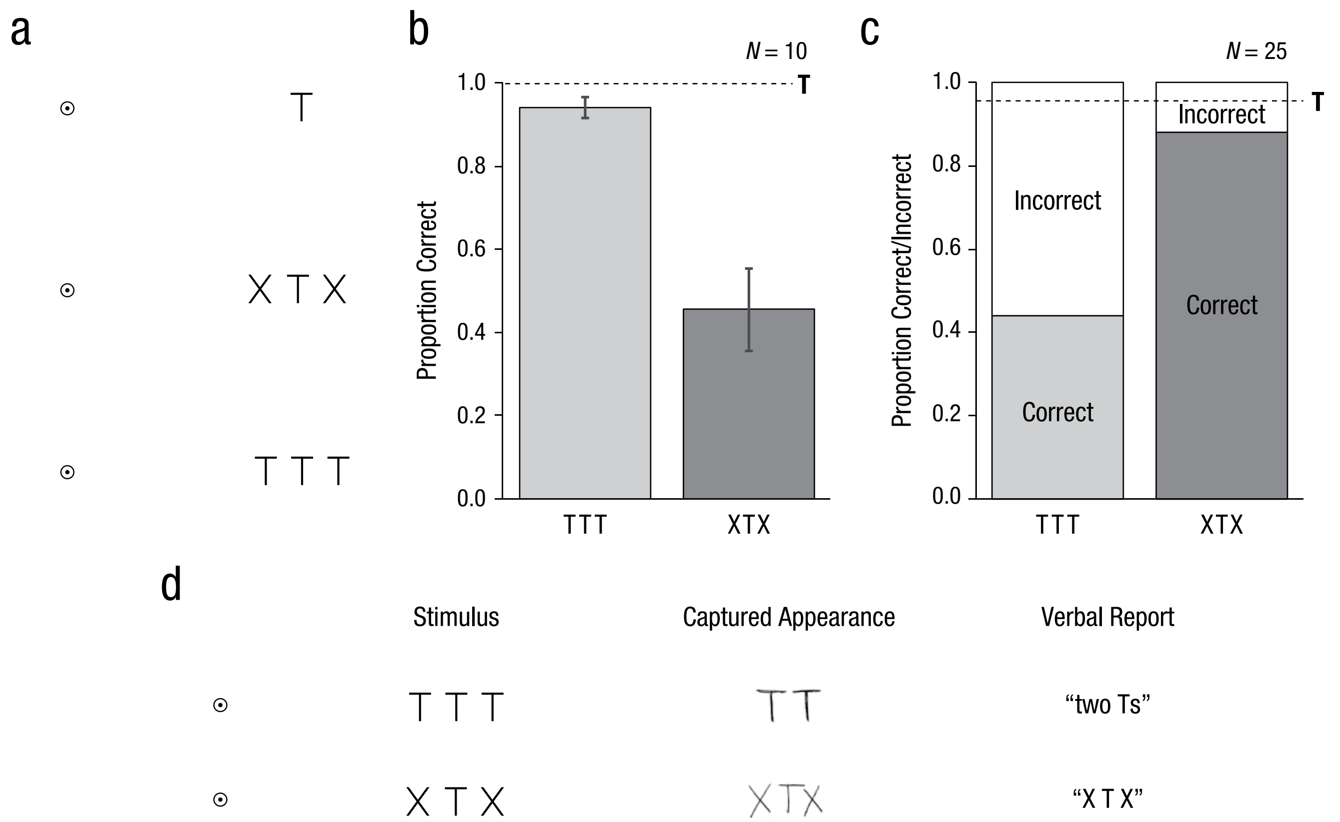

Fig. 1. Illustration of crowding and identity crowding, results, and demonstration of redundancy masking. In the crowding and identity-crowding paradigm (a), when observers are asked to fixate on the dot (left) and identify the target letter (T) presented in the periphery (right), most observers are able to identify the T when it is unflanked (top). Identification is usually more difficult when the target is flanked by other letters (middle). In identity crowding, the target is flanked by identical items (bottom). Similar stimuli were used in Experiments 1 and 2. The graphs show the results of (b) Experiment 1 and (c) Experiment 2. For Experiment 1, the proportion of correct responses is shown for the two flanker conditions. For Experiment 2, the proportion of correct and incorrect responses is shown for the two flanker conditions. The dashed lines show the proportion correct for the unflanked target (T). Error bars in (b) indicate standard errors of the mean. Sample responses in Experiment 2 are shown in (d). When three Ts were presented, most participants reported (and drew) two Ts, illustrating redundancy masking. When Xs flanked the target, most participants correctly reported three items and the target as a T. Two representative participant drawings are shown here for each condition.

almost perfect performance in identity crowding (Experiment 1). Next, to overcome the aforementioned challenges, we used an unconstrained free-report-and-drawing paradigm with gaze-contingent stimulus presentation (Experiment 2). Participants frequently reported only two instead of the three presented identical letters (i.e., performance was poor). Our results reveal a new effect that we call *redundancy masking*, in which the number of perceived items is reduced.

Method

Participants

In Experiment 1, 10 paid students participated (5 female, 5 male; mean age = 23.1 years). In Experiment 2, 25 students participated for course credit (16 female, 9 male; mean age = 26.0 years). The sample sizes were based on studies using similar methodologies, with a significant increase in the number of participants in Experiment 2 to compensate for the comparably small number of trials (Sayim & Wagemans, 2017). All participants had normal or corrected-to-normal visual acuity.

Apparatus and stimuli

Stimuli were presented on a CRT monitor (HP P1230 with a refresh rate of 110 Hz in Experiment 1; Sony Trinitron GDM-F520 with a refresh rate of 120 Hz in Experiment 2; resolution: 1,152 × 864 pixels). A head-and-chin rest was used to stabilize the head position. Participants viewed the monitor from a distance of 57 cm. The main target stimulus consisted of the letter T, presented at 10° eccentricity. In three separate conditions, the target was presented alone, flanked by two Xs (XTX), or flanked by two Ts (TTT; Fig. 1a). In

Experiment 1, the letters E, F, H, K, L, N, V, X, and Z were used as additional targets (see the Procedure section). All letters were in Microsoft Yi Baiti font (redrawn in Experiment 2). The letters were 1.4° high and 1.1° wide (with small deviations depending on the letters in Experiment 1). The center-to-center spacing between the target and each flanker was 1.3°. A fixation dot was presented in the center of the screen. All elements were black with a luminance of 0.48 cd/m^2 in Experiment 1 and 0.1 cd/m^2 in Experiment 2, presented on a gray background (50.1 cd/m^2 in Experiment 1; 50.5 cd/m^2 in Experiment 2). In Experiment 2, participants' gaze was tracked with an EyeLink 1000 eye tracker (SR Research, Kanata, Ontario, Canada). In this experiment, a drawing board was positioned in front of the head-and-chin rest. Drawings were made on paper with a standard pen. Verbal reports were recorded by the experimenter.

Procedure

In Experiment 1, stimuli were presented for 150 ms, randomly to the left or right of fixation. Participants were informed that three letters would be presented, and they were instructed to indicate the central letter by pressing the corresponding key on a keyboard. Participants completed 10 blocks with 100 trials each. Each letter (E, F, H, K, L, N, T, V, X, Z) was presented 10 times per block. In 8 blocks, the target was flanked in random order by Xs in half of the trials and Ts in the other half. There were two conditions of interest. Normal crowding, using the XTX stimulus, and identity crowding, using the TTT stimulus. Each block contained the main target stimuli of XTX and TTT 5 times; hence, each was presented 40 times in total. In the remaining 2 blocks, unflanked performance was measured (20 trials per target letter). Note that the non-T target letters were used only as filler stimuli so that we could measure performance on the main targets (XTX and TTT) without obvious repetitions.

In Experiment 2, each participant completed one trial with the XTX stimulus, TTT stimulus, and T stimulus, respectively. Stimuli were presented in the right visual field at the same eccentricity as in Experiment 1 (10°). We used eye tracking to present the stimuli only when participants kept central fixation. Viewing time was unconstrained. Participants were asked to freely draw and verbally report what they saw without any constraints. Crucially, no instructions were given that allowed participants to infer that three letters were present, unlike in Experiment 1. The drawings were made at the center of the drawing board, approximately aligned with fixation, requiring eye movements along the vertical to alternate between looking at the screen and the drawings. Half of the participants started with the XTX condition, and the other half started with the TTT condition. The unflanked target was always presented last. The verbal response was classified as correct if it fulfilled two criteria: Participants reported that there was a central letter (requiring that three items were reported) and that the letter was a T. The drawings were made to avoid reliance on a single measure (i.e., the free verbal reports) and to get a good understanding of how the stimuli appeared to the participants. Before we ran each experiment, participants performed a number of training trials to get familiarized with the method. In Experiment 1, the training stimuli were randomly selected from the stimulus set. In Experiment 2, the training stimuli consisted of the same elements as the target and the flankers, arranged in abstract geometric configurations.

Results

In Experiment 1, the proportion of participants who correctly reported T in the identity-crowding condition (TTT) was high (.94, *SE* = .03; Fig. 1b). In the normal-crowding condition (XTX), performance was clearly worse (proportion correct = .46, *SE* = .10), $t(9) = 5.60$, $p < .001$, Cohen's $d = 2.15$. Proportion correct for the unflanked T was 1.0. The proportion of participants who erroneously reported a flanker (X) was .33 (*SE* = .04) in the XTX condition. Importantly, the flanker report rate could not be determined in the TTT condition. The average proportion correct for the other target letters was .62 (*SE* = .06) with X flankers and .82 (*SE* = .04) with T flankers (proportion correct for the unflanked stimulus = .98, *SE* = .004). This result seems to support the hypothesis that crowding is comparatively weak when all items are the same. However, the use of a standard crowding paradigm to measure performance when the target and the flankers are identical has, as outlined above, several shortcomings relating to task demands, prior knowledge, and the fact that report of a flanker is counted as "correct" (see also Sayim & Cavanagh, 2013). We addressed these concerns in Experiment 2.

The results of Experiment 2 showed that targets were not reported more accurately in identity crowding compared with normal crowding (Fig. 1c). On the contrary, the proportion correct in the free verbal report was lower in the identity-crowding condition (.44) than in the normal-crowding condition (.88; odds ratio = 0.107, Fisher's exact test, $p < .005$). Most remarkably, all errors in the identity-crowding condition were due to missing one of the three items, reporting two Ts instead of three. Participants' drawings matched their free verbal responses, confirming that they perceived two Ts rather than three in the identity-crowding condition (Fig. 1d). Hence, the perceived number of items in the identity-crowding condition was lower than the number of presented items, revealing a strong case of diminishment (Coates, Wagemans, & Sayim, 2017; Sayim & Wagemans, 2017). Notably, 96% of the responses in the identity-crowding condition contained the letter T, and 92% contained no other letter than T. Hence, it is not surprising that standard identification tasks, as in Experiment 1, result in "correct" responses (reporting the letter T) and thereby miss the pronounced misperception of the total number of items (two Ts instead of three).

The rate of correct responses in the normal-crowding condition of Experiment 2 was relatively high, compared with that in Experiment 1, presumably because of long presentation times (Styles & Allport, 1986) and multiple views of the same stimulus (Sayim & Wagemans, 2017). Remarkably, accuracy in the identity-crowding condition was nevertheless very poor, suggesting that redundancy masking is strong even under conditions that benefit performance in normal crowding.

In an additional experiment (Experiment 3), we used printouts of the XTX and TTT drawings from Experiment 2 and asked 100 naive participants (4 participants per drawing; 61 female; mean age = 23.8 years) to indicate the central—or hypothetically central—target letter (for representative drawings, see Fig. 1d). In the identity-crowding condition (TTT), 84% of the participants responded that the target letter was a T (SE = 5%), and in the normal-crowding condition (XTX), 90% of the participants responded that the target letter was a T (SE = 5%). Hence, even when there were only two Ts in a drawing (and therefore no central T), participants mostly reported the letter T. This result supports the finding of Experiment 1. When participants were asked to report the central of three letters and they saw only two Ts, the best response (or guess) was still that it was a T.

Overall, the results show that stimuli in identity crowding were not perceived better than in normal crowding. Rather, a remarkable and highly consistent error characterized identity-crowded appearance—only two instead of three Ts were reported by the majority of participants (Experiment 2; see also Fig. 1d). This type of diminishment error cannot be captured with a standard crowding task like the one used in Experiment 1. Using the drawings of Experiment 2 as representations of stimulus appearance and asking naive participants to report the (hypothetical) central target letter, we confirmed that correct responses are very likely in identity crowding even when only two items are perceived.

Discussion

These results demonstrate a strong diminishment effect in crowded displays (Sayim & Wagemans, 2017). Unlike in normal crowding, stimuli in identity crowding are characterized by maximum target–flanker similarity, high regularity, and redundancy, which, we suggest, yields a new type of error through a mechanism that we call redundancy masking. Poor performance in identity crowding is mainly caused by the "disappearance" or masking of an entire item (Tye, 2014), instead of the perceived "jumble" that is seen in normal crowding.

Our results provide strong evidence against the hypothesis that targets in identity crowding are identified better than in normal crowding (Block, 2012). Conversely, they support the hypothesis that target disruption is stronger in identity crowding than in normal crowding (Taylor & Sayim, 2018).

The unconstrained free-report paradigm is crucial to revealing this new effect because standard forced-choice methods, as in Experiment 1, conflate cases of genuinely perceiving the central target and mistaking three for two letters. By contrast, in Experiment 2, participants were allowed to report the number of letters and their identity, thereby providing insight into unbiased stimulus appearance. The results of Experiment 3, with a high rate of correct target identifications in drawings containing only two letters, support the view that participants will report a central T when all they really see is two Ts and that this may underlie the seemingly better performance in identity crowding (Taylor, 2013).

Redundancy masking shares characteristics with crowding, masking, and statistical summary representations. Regarding crowding, our findings are at odds with the assumption that it hinders only feature integration and not feature detection (Pelli, Palomares, & Majaj, 2004). Although we did not use a classic detection task, our results show the perceived absence of one of the items akin to a "miss" in masking paradigms. However, the temporal and spatial features of our stimuli diverged from those used in traditional masking studies (Breitmeyer & Öğmen, 2006). Although statistical summary representations may occur for as few as two items, they are usually assumed to be effective when larger numbers of items are displayed (Whitney & Yamanashi Leib, 2018). A limit of attentional resolution (He, Cavanagh, & Intriligator, 1996) may play a role in redundancy masking, but the failure to detect all three items is not predicted by this account. Still unknown are the underlying mechanisms of redundancy masking, whether items lost by redundancy masking still prime observers (Yeh, He, & Cavanagh, 2012) or bias observers (Kouider, Berthet, & Faivre, 2011; Manassi & Whitney, 2018), and whether redundant elements are lost in normal crowding. By revealing unbiased visual appearance, our findings demonstrate a remarkably strong illusion with crowded stimuli, suggesting a mechanism that reduces the perceived number of redundant elements.

Action Editor

Edward S. Awh served as action editor for this article.

Author Contributions

Both authors developed the study concept and contributed to the study design. Testing, data collection, and data analysis were performed by B. Sayim. Both authors interpreted the results. B. Sayim drafted the manuscript, and H. Taylor provided critical revisions. Both authors approved the final manuscript for submission.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This research was supported by the Swiss National Science Foundation (PP00P1_163723) and an Early Career Fellowship from the Leverhulme Trust and the University of Birmingham (ECF-2015-088).

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797619847166

Open Practices

The drawing results are available in the Supplemental Material available online; further data and materials are available from the authors on request. The design and analysis plans were not preregistered.

References

Block, N. (2012). Seeing and windows of integration. *Thought*, *2*, 29–39.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177–178.

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

Coates, D. R., Wagemans, J., & Sayim, B. (2017). Diagnosing the periphery: Using the Rey–Osterrieth complex figure drawing test to characterize peripheral visual function. *i-Perception*, *8*(3). doi:10.1177/2041669517705447

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, *383*, 334–337.

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

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision*, *8*, 255–279.

Kouider, S., Berthet, V., & Faivre, N. (2011). Preference is biased by crowded facial expressions. *Psychological Science*, *22*, 184–189.

Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research*, *48*, 635–654.

Manassi, M., & Whitney, D. (2018). Multi-level crowding and the paradox of object recognition in clutter. *Current Biology*, *28*, 127–133.

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

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

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

Sayim, B., Westheimer, G., & Herzog, M. H. (2010). Gestalt factors modulate basic spatial vision. *Psychological Science*, *21*, 641–644.

Styles, E. A., & Allport, D. A. (1986). Perceptual integration of identity, location and colour. *Psychological Research*, *48*, 189–200.

Taylor, H., & Sayim, B. (2018). Crowding, attention and consciousness: In support of the inference hypothesis. *Mind & Language*, *33*, 17–33.

Taylor, J. H. (2013). Is the grain of vision finer than the grain of attention? Response to Block. *Thought*, *2*, 20–28.

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, *32*, 1349–1357.

Tye, M. (2014). Does conscious seeing have a finer grain than attention? *Thought*, *3*, 154–158.

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology*, *69*, 105–129.

Yeh, S., He, S., & Cavanagh, P. (2012). Semantic priming from crowded words. *Psychological Science*, *23*, 608–616.