Letters lost

Sayim, Bilge; Taylor, John

DOI:
10.1177/0956797619847166

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

Bilge Sayim[1,2*] & Henry Taylor[3]

[1]Institute of Psychology, University of Bern, Bern, Switzerland
[2]Univ. Lille, CNRS, UMR 9193 - SCALab - Sciences Cognitives et Sciences Affectives, F-59000 Lille, France
[3]Department of Philosophy, University of Birmingham, Birmingham, UK
*Corresponding author: bilge.sayim@univ-lille.fr

Author's accepted manuscript
Forthcoming in *Psychological Science* (DOI not yet available)
Please Cite Published version

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 susceptible to these influences, and next - to minimize them - an unconstrained full report and drawing paradigm. Three letters were presented in the periphery. In Experiment 1, ten observers were asked to identify the central target letter. In Experiment 2, 25 observers freely named and drew what they saw. When three *identical* letters were presented, performance was almost perfect in Experiment 1, but very poor in Experiment 2 where most observers reported only two letters. Our study reveals limitations of standard crowding paradigms, and it uncovers a hitherto unrecognised effect we call "redundancy masking".

Introduction

We usually have the mistaken impression of unconstrained, high resolution access to the 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, letter identification deteriorates when the target is surrounded by flanking letters (Fig. 1a). Crowding is generally stronger when the target and the flankers are nearby (Toet &

Levi, 1992), similar (Kooi et al., 1992), and group together (Herzog et al., 2015; Sayim et al., 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 target and flankers are similar, so we might expect that target identification in identity-crowding is difficult. On the other, it was recently proposed that target identification in identity-crowding is *superior* to normal crowding (Block, 2012; cf. Taylor & Sayim, 2018). To evaluate these two hypotheses, an experimental paradigm is needed that can test what is genuinely seen in (identity-) crowding.

Identity-crowding has unique methodological challenges. Since the target and the flankers are the same, it is difficult to separate target 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 almost perfect performance in identity crowding. Next, to overcome the aforementioned challenges, we used an unconstrained full report and drawing paradigm with gaze-contingent stimulus presentation. Observers frequently reported only two instead of the three presented identical letters, i.e., performance was poor. Our results reveal a new effect we call “redundancy masking”, in which the number of perceived items is reduced.

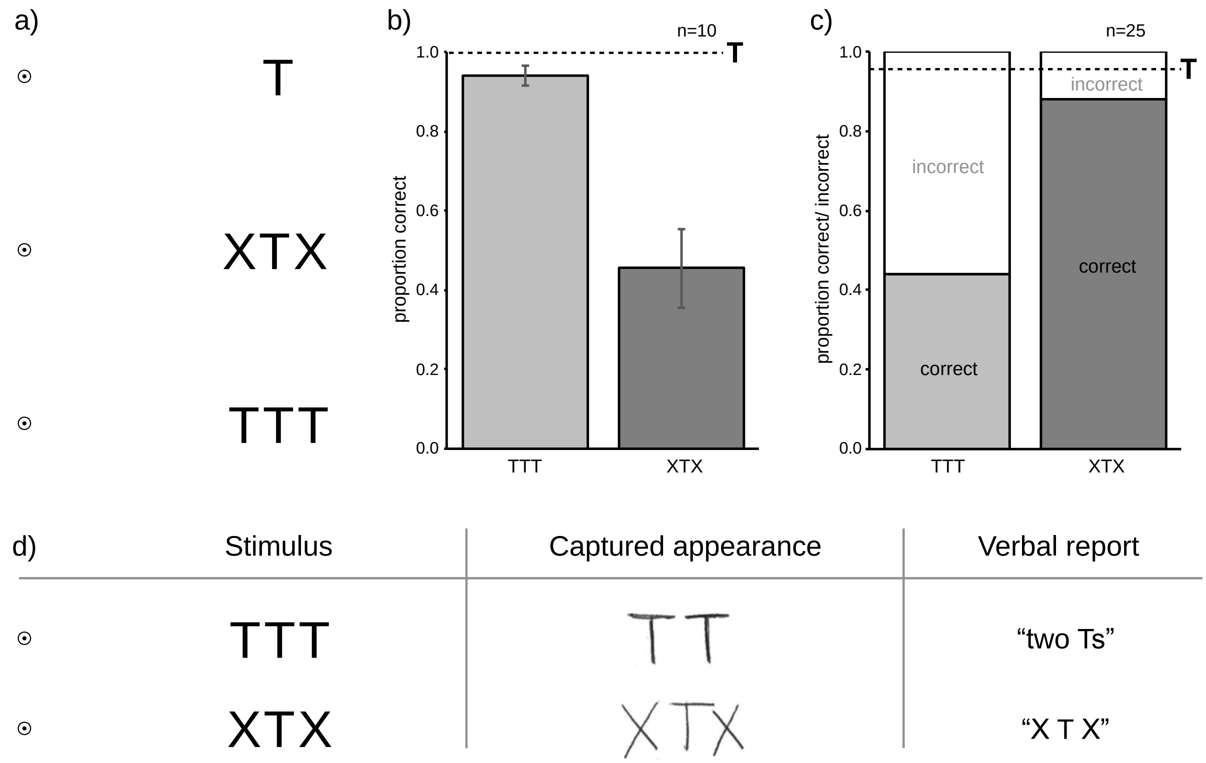

Figure 1: **a)** When fixating the upper disc, most observers are able to identify the T on the right. Identification is more difficult when the target is flanked by letters (middle). In ‘identity-crowding’, the target is flanked by identical items (bottom). **b)** Results of Experiment 1. Proportion correct was higher when the target and the flankers were the same (TTT) compared to when they were different (XTX). The dashed line shows unflanked proportion correct. Error bars indicate standard errors of the mean. **c)** Results of Experiment 2. Proportion correct was lower when the flankers were the same as the target (TTT) compared to when they were different (XTX), the opposite results of Experiment 1. **d)** Illustration of ‘redundancy masking’. Three Ts presented in the periphery appeared and were reported as two Ts. When Xs flanked the target, no redundancy masking occurred. Two representative drawing results are shown under “Captured appearance”.

Methods

Participants

In Experiment 1, ten paid students participated (5 female, 5 male; mean age = 23.1). In Experiment 2, 25 students participated for course credit (16 female, 9 male; mean age = 26.0). The sample sizes were based on studies using similar methodologies, with a significant increase of 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 Exp. 1, and Sony Trinitron GDM-F520 with a refresh rate of 120 Hz in Exp 2; resolution: 1152 x 864). 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 degrees eccentricity. In three 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, Z were used as additional targets (see procedure). All letters were of Microsoft Yi Baiti font (redrawn in Exp. 2). The letters were 1.4 degrees high and 1.1 degrees 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 degrees. A fixation dot was presented in the center of the screen. All elements were black with a luminance of 0.48 cd/m^2 (0.1 cd/m^2, in Experiment 2) presented on a gray background (50.1 cd/m^2; 50.5 cd/m^2 in Experiment 2). In Experiment 2, observers' gaze was tracked with an EyeLink 1000 (SR Research). A drawing board was positioned in front of the head/ 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. Subjects were informed that three letters were presented and were instructed to indicate the central letter by pressing the corresponding key on a keyboard. Observers completed 10 blocks with 100 trials. Each letter (E, F, H, K, L, N, T, V, X, Z) was presented 10 times per block. In eight 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 5 times the main target stimuli XTX and TTT, hence, each was presented 40 times in total. In the remaining two blocks, unflanked performance was measured (20 trials per target letter). Note that the non-T target letters were only used as filler stimuli to be able to measure performance on the main targets (XTX and TTT) without obvious repetitions.

In Experiment 2, each participant completed one trial with the XTX, TTT, 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. Observers were asked to draw with free viewing, and verbally report what they saw without any constraints. Crucially (unlike in Experiment 1) no instructions were given that allowed subjects to infer that three letters were present. 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, the other half with the TTT condition. The unflanked target was always presented last. The verbal response was classified as correct if it fulfilled two criteria: subjects reported that there was a central letter (requiring that three items were reported), and that it 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 subjects. Before 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, they consisted of the same elements as the target and the flankers, arranged in abstract geometric configurations.

Results

In Experiment 1, the proportion of correctly reporting "T" in the identity-crowding condition (TTT) was high (0.94, SE=0.03; Fig. 1b). In the normal crowding condition (XTX), performance was clearly worse (proportion correct=0.46, SE=0.10; t-test: $t(9)= 5.60$, $p<0.001$; Cohen's $d=2.15$). Proportion correct for the unflanked T was 1. The proportion of erroneously reporting a flanker (X) was 0.33 (SE=0.04) in the XTX condition. Importantly, the flanker report rate cannot be determined in the TTT condition. The average proportion correct for the other target letters was 0.62 (SE=0.06) with X-flankers, and 0.82 (SE=0.04) with T-flankers (unflanked proportion correct was 0.98; SE=0.004). 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 to do with task demands, prior knowledge and the fact that report of a flanker is counted as 'correct' (see also Sayim & Cavanagh, 2013). We addressed these in Experiment 2.

The results of Experiment 2 showed that targets were not reported more accurately in identity- compared to normal crowding (Fig. 1c). To the contrary, proportion correct in the free verbal report was lower in identity-crowding (0.44) compared to normal crowding (0.88; Odds-Ratio=0.107, Fisher's Exact Test, $p<0.005$). Most remarkably, all errors in the identity-crowding condition were due to missing one of the three items, reporting two Ts instead of three. The 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 by crowding (Coates, Wagemans, & Sayim, 2017; Sayim & Wagemans, 2017). We call this effect 'redundancy masking' – a 'redundant' item (the T) is not (consciously) perceived, or 'masked'. Notably, 96% of the responses in the identity-crowding condition contained the letter 'T' and 92% 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 T's instead of three).

Compared to Experiment 1, the rate of correct responses in the normal crowding condition of Experiment 2 was relatively high, presumably due to long presentation times (Styles and 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 (see below) 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 naïve participants (four participants per drawing; 61 female, mean age = 23.8) to indicate what was the central - or hypothetically central – target letter (Fig. 1d shows two representative drawings). In the identity-crowding condition (TTT), 84% (SE=0.05), and in the normal crowding condition (XTX), 90% (SE=0.05) of the participants responded that the target letter was a T. 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 asked to report the central of three letters, and participants only see two Ts, the best response (or guess) is 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 as in Experiment 1. Using the drawings of Experiment 2 as representations of stimulus appearance, and asking naïve participants to report the (hypothetical) central target letter, 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 crowding (Sayim & Wagemans, 2017). Unlike 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 we call 'redundancy masking'. Instead of the perceived 'jumble' that is seen in normal crowding, poor performance in identity-crowding is mainly caused by the 'disappearance' or masking of an entire item (Tye, 2014).

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- than in normal crowding (Taylor & Sayim 2018).

The unconstrained free-report paradigm is crucial to revealing this new effect as 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 result of Experiment 3, with a high rate of 'correct' target identifications in drawings containing only two letters, supports the view that subjects 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 only hinders feature integration and not feature detection (Pelli et al., 2004). While 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 diverge from those used in traditional masking studies (Breitmeyer & Öğmen, 2007). 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 et al., 1996), may play a role in redundancy masking, but the failure to detect all of three items is not predicted by this account. What are the underlying mechanisms of redundancy masking, whether items lost by redundancy masking still prime (Yeh et al., 2012) or bias observers (Kouider et al., 2011; Manassi & Whitney, 2018), and whether redundant elements are lost also in normal crowding, are open questions. 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.

Declaration of Conflicting Interests

The authors declare no conflicts of interest with respect to their authorship and 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).

References

Block, N. (2012). Seeing and windows of integration. Thought, 2(1), 29–39.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. Nature, 226(5241), 177–178.

Breitmeyer, B. G., & Öğmen, H. (2006). Visual Masking: Time Slices Through Conscious and Unconscious Vision. Oxford: Oxford University Press.

Coates, D. R., Wagemans, J., Sayim, B. (2017). Diagnosing the visual periphery: Using the Rey-Osterrieth Complex Figure Test to evaluate peripheral visual function. i-Perception. doi: 10.1177/2041669517705447.

He, S., Cavanagh, P. & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. Nature, 383(6598), 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) 5, 1-18.

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(2), 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(5), 635–654.

Manassi, M., Whitney, D. (2018). Multi-Level Crowding and the Paradox of Object Recognition in Clutter. Current Biology 28(3), 127-133.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. Nature Neuroscience, 4(7), 739–744.

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): e71188.

Sayim, B., & Wagemans, J. (2017). Appearance changes and error characteristics in crowding revealed by drawings. Journal of Vision, 17(11):8, 1–16.

Sayim, B., Westheimer, G., Herzog, M. H. (2010). Gestalt Factors Modulate Basic Spatial Vision. *Psychological Science*, 21(5), 641-644.

Taylor, H., & Sayim, B. (2018). Crowding, attention and consciousness: In support of the inference hypothesis. Mind Lang.;1–17.

Taylor, H. (2013). Is the grain of vision finer than the grain of attention? Response to

Block. Thought 2(1), 20–28.
Tye, M. (2014). Does conscious seeing have a finer grain than attention? Thought, 3(2), 154–158.
Whitney, D., & Yamanashi Leib, A. (2018). Ensemble Perception. Annual Review of Psychology, Vol. 69:105-129.
Yeh, S., He, S. & Cavanagh, P. (2012). Semantic priming from crowded words. Psychological Science, 23(6), 608–616.