Journal of Memory and Language 145 (2025) 104689

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

A lifetime of reading experience facilitates the perception of crowded letters

Kurt Winsler *, Steven J. Luck

Center for Mind & Brain and Department of Psychology, University of California – Davis, USA

ARTICLE INFO

Keywords:
Letter perception
Reading
Perceptual learning
Visual crowding
Visual experience
Visual working memory

ABSTRACT

Visual perception is ordinarily impaired for objects that are tightly crowded by other objects. This might be expected to make reading very difficult given that letters are tightly crowded together within words. However, a lifetime of reading experience may lead to changes in visual processing that reduce the effects of crowding on letters. Study 1 examined this hypothesis experimentally by comparing crowding thresholds (measured as the closest spacing that yields recognition accuracy of 82% correct) for upright letters, inverted letters, and Gabor patches in 60 experienced readers of English. We found that crowding thresholds were reduced for upright letters compared to other stimuli classes, especially for stimuli close to the fovea. In other words, experienced readers could tolerate closer spacing for highly familiar upright letters than for less familiar types of stimuli. Crowding thresholds were also reduced to the right of fixation, matching the left-to-right direction of English reading. Study 2 measured crowding in 250 observers and asked whether individual differences in proxies of reading experience were associated with reduced crowding. We found that higher scores on these proxy measures were associated with lower crowding thresholds for upright letters, especially in the right visual field. These results provide evidence that a lifetime of reading experience alters aspects of visual perception, such that upright letters can be perceived under more-crowded conditions than other stimuli. At a practical level, this means that deleterious effects of letter crowding are significantly reduced for experienced readers, which has implications for both models of visual crowding and reading.

Introduction

Reading is a demanding task for the visual system, requiring fast-paced eye movements and efficient recognition of visually complex words from sets of crowded letters, all while integrating this information into dynamic semantic and syntactic structures. Although reading is a challenging task for the untrained visual system, it is a common part of daily life for many humans and becomes effortless for many readers after years of practice, leading to average reading rates of about 4 words per second (Brysbaert, 2019). Reading is a behavior learned through experience, and was culturally invented too recently to have driven significant brain evolution. Hence, reading offers a unique avenue to uncover effects of real-world experience on vision.

In addition, reading is much more constrained than most common visual tasks. For example, proficient readers of the Latin script have viewed a relatively small set of letter shapes millions of times, in a particular part of the visual field and attentional context (scanning across the horizontal meridian), and for a particular purpose (to bind letters into words, and ultimately into meaning). This highly constrained and frequently repeated task may lead to systematic changes in aspects of vision that adapt visual processing to the particular requirements of reading through neuroplasticity. It is clear that relatively short-term laboratory experiences can impact low-level visual processing (Sagi, 2011), but the effects of long-term real-world experience with specific stimuli such as letters through particular behaviors like reading on relatively low-level visual phenomena are largely unknown.

One visual phenomenon that is particularly relevant to reading is visual crowding. In a classic study that documented a fundamental limit on visual perception, Bouma (1970) demonstrated that a peripheral letter could be accurately identified when presented in isolation but not when flanked by nearby letters. This reduction in recognition due to nearby items is known as visual crowding (Fig. 1 provides a demonstration). Importantly, Bouma also demonstrated that the distance at which flankers crowd a target scales linearly with the distance of that target from fixation (which has subsequently been termed "Bouma's law"; Pelli & Tillman, 2008). That is, a target letter presented 4° from

* Corresponding author at: Center for Mind & Brain, 267 Cousteau Pl, Davis, CA 95616, USA.
E-mail address: kpwinsler@ucdavis.edu (K. Winsler).

https://doi.org/10.1016/j.jml.2025.104689
Received 28 January 2025; Received in revised form 24 June 2025; Accepted 9 August 2025
Available online 13 August 2025

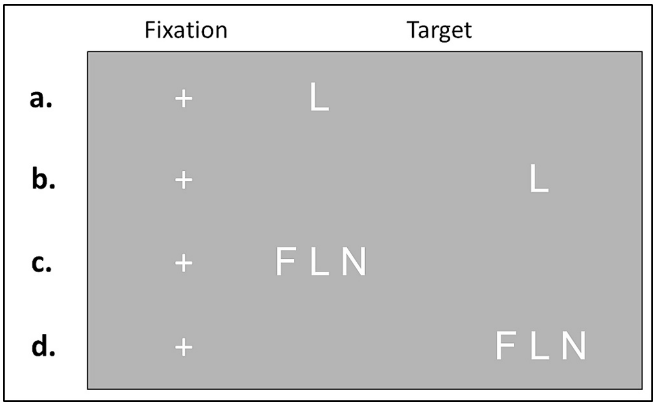

Fig. 1. Demonstration of crowding. Keep your eyes on the central fixation crosshair for a given row while trying to perceive the letter(s) for that row. The letter L can be perceived quite well at both a near eccentricity (**a**) and a far eccentricity (**b**). However, the L becomes much more difficult to perceive when crowded by nearby flanker letters at the near eccentricity (**c**) and practically impossible to perceive at the far eccentricity (**d**). The fact that the L is easily perceived in isolation indicates that low-level visual acuity is not the reason for the difficulty of perceiving the L when it is crowded.

fixation would be degraded by flanking letters 2° on each side to approximately the same extent that a target letter presented 10° from fixation would be degraded by flanking letters 5° on each side.

Visual crowding is now known to occur for many types of stimuli and is a major source of functional limitation in real-world vision. It reflects fundamental properties of the visual system and provides clues to the nature of object recognition and visual representation (Levi, 2008; Whitney & Levi, 2011; Manassi & Whitney, 2018). Bouma's original study and much subsequent research (Pelli et al., 2007; Chung, 2007; He, Legge, & Yu, 2013) has been concerned with crowding as a limitation on reading, because letters are highly crowded within words. Indeed, readers use crowded information from upcoming letters and words that are as many as 15 character spaces away from fixation (Rayner, 2009), and there is now considerable evidence that crowding is a key limiting factor in both typical reading (Risse, 2014; Frömer et al., 2015; He & Legge, 2017) and disorders of reading such as dyslexia (Joo et al., 2018).

Here, however, we invert this relationship between crowding and reading to ask a different question: To what extent does our massive experience with reading reduce the effects of crowding? We focus on three aspects of reading that might lead to corresponding changes in crowding. First, reading involves massive experience with a relatively small number of visually-similar characters, which are almost always read in an upright orientation and frequently occur in closely spaced strings (i.e., words). Further, both the precise identity and relative position of letters in words carry important information which disambiguate word identities during reading (e.g., the strings "salt" and "slat" have completely different meanings), though transposed letter effects do reveal that in some cases words with swapped letter positions are recognizable (e.g., Perea & Lupker, 2004). This massive but highly constrained reading experience may cause changes in how the visual system processes letters, such that crowding is minimized for upright characters. For example, letters could be processed using a narrower spatial zone of integration or could be processed with reduced lateral inhibition. Alternatively, neurons may become more narrowly tuned to specific letters, such that different letters are more discriminable and can therefore be recognized even when degraded by information from nearby stimuli.

A second relevant aspect of reading is that the crowding of letters must be overcome primarily in the fovea and parafovea, whereas many other types of natural vision do not require individuating a large number of discrete elements that are so closely spaced near the fovea. As a result, crowding for upright letters may be particularly reduced near the point of fixation.

A third relevant aspect of reading is that readers of scripts with a left-to-right organization, such as English and Spanish, rely heavily on parafoveal information from the right side of fixation to control their eye movements (Schotter, Angele, & Rayner, 2012), to guide recognition of currently fixated words (Veldre & Andrews, 2016), and in some cases to recognize upcoming words without fixating them (Rayner et al., 2011). Consequently, readers of such languages may develop mechanisms that specifically minimize crowding to the right of fixation (e.g., attentional mechanisms that are preferentially applied to stimuli to the right of fixation).

There is already some evidence that crowding effects differ between letters and other types of stimuli. For example, recognition accuracy is better for crowded letters than for crowded symbols (Grainger, Tydgat, & Issele, 2010), and this advantage has been shown to be true particularly for horizontally oriented arrays rather than vertically oriented arrays (Vejnović & Zdravković, 2015). However, these experiments compared letters with symbols or digits (e.g., Tydgat and Grainger, 2009; Grainger, Tydgat, & Issele, 2010; Vejnović & Zdravković, 2015; Winsler, Grainger, & Holcomb, 2022), making it possible that the differences in accuracy may be due to differences in the physical stimuli themselves rather than differences in experience with the stimuli. Additionally, almost all of these studies used behavioral accuracy at a single target-flanker spacing as the dependent measure, potentially reflecting post-perceptual processes (e.g. higher recognition accuracy regardless of flanker spacing) rather than crowding itself (but see Experiment 5 in Grainger, Tydgat & Issele, 2010). Moreover, no prior research has examined whether individual differences in reading experience are associated with individual differences in crowding. Thus, prior research does not clearly address the question of whether massive visual experience with reading impacts crowding.

There are multiple underlying mechanisms that could potentially lead to reduced crowding for upright letters, including changes in the mechanisms responsible for crowding per se and improvements in letter discriminability that allow letters to be accurately identified even in the presence of nearby flankers. In either case, the *functional* consequence would be a reduction in crowding for upright letters, which would presumably minimize the deleterious consequences of crowding on reading, allowing for faster and more accurate reading. This would indicate that massive visual experience with a class of stimuli can reduce the deleterious consequences of crowding for that stimulus class.

The present studies

The present study addressed the question of whether reading experience impacts the effects of crowding using two preregistered, converging approaches: a) an experimental approach in which we compared crowding for upright letters with crowding for less familiar stimuli, and b) an individual-differences approach in which we asked whether specific attributes of crowding are related to proxy measures of lifetime reading experience. As illustrated in Fig. 2, both experiments used multiple target-flanker distances to establish the target-flanker distance at which crowding occurred. We refer to this as the *crowding threshold* (also known as critical distance), defined as the target-flanker distance that yielded a target discrimination accuracy of 82 % correct. This allowed us to determine whether familiarity impacts crowding per se rather than impacting overall accuracy. We also tested multiple target eccentricities on both the left and right sides of fixation so that we could determine whether the spatial profile of crowding is different for upright letters (i.e., decreased crowding for upright letters near fixation and in the right visual field). Note that we focused here on radial crowding along the horizontal meridian, which is the primary type of crowding experienced during reading of left-to-right scripts.

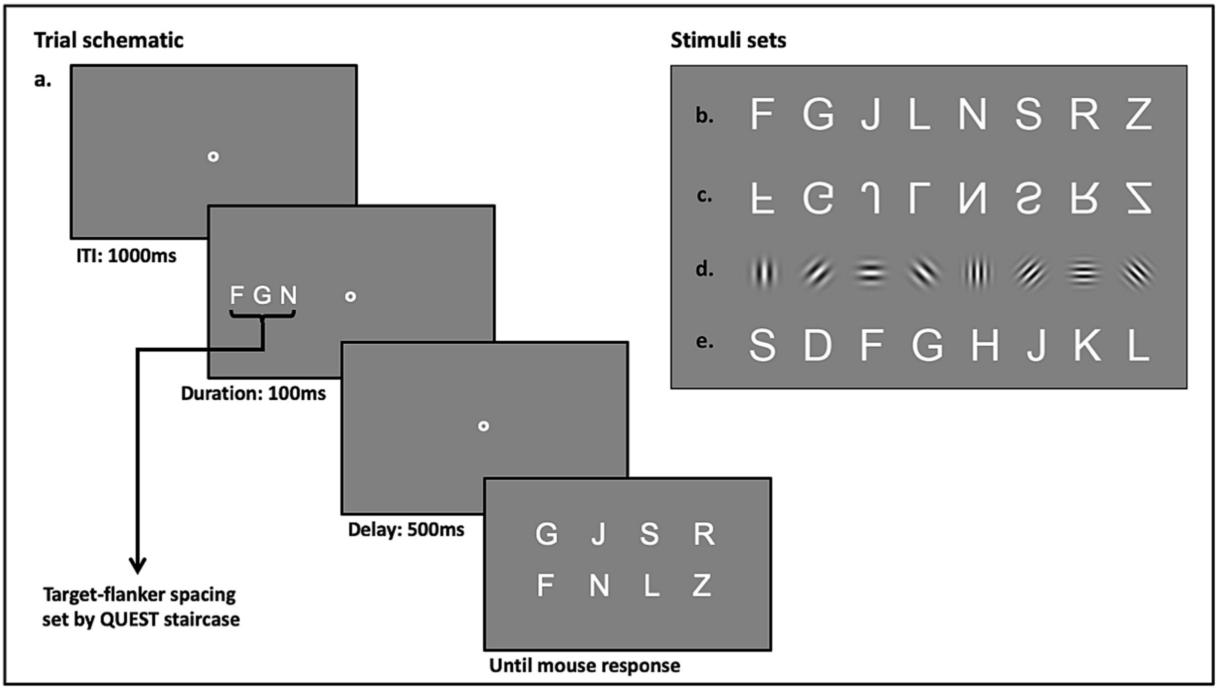

Fig. 2. **a.** Example of a trial with upright letters in Study 1. The target was the central item in each three-item string of letters. After a period of fixation on the central fixation point, a target and two flankers appeared for 100 ms. The target was unpredictably placed at one of three eccentricities in the left or right visual field. After a 500 ms of delay, an array of eight choice letters appeared and the participant used the mouse to click on the letter that matched the target. We also included catch trials on which the target was presented without any flankers. **b.** Upright letter stimuli used in Studies 1a and 1b. **c.** Inverted letter stimuli used in Study 1a. **d.** Gabor patch stimuli used in Study 1b. **e.** Upright letter stimuli used in Study 2.

Study 1: Experimental manipulation of stimuli

In Study 1a, we assessed the effect of experience by comparing crowding for upright letters with crowding for inverted letters. Inverted letters are much less familiar than upright letters, but they provide a much better control for low-level stimulus properties than other stimuli such as symbols. For example, inverting a letter does not change perimetric complexity (Pelli et al., 2006), spatial frequency content, or basic visual features. This makes it possible to be nearly certain that any differences in crowding between upright and inverted letters were caused by differences in experience rather than by intrinsic differences in visual properties. We tested three predictions: a) that the visual system can tolerate closer target-flanker distances for upright letters than for inverted letters; b) that the advantage for upright letters would be particularly strong near fixation; and c) that this advantage would be greater to the right of fixation than to the left. Study 1b tested the same predictions by comparing upright letters with Gabor patches, which are unfamiliar but closely match the tuning properties of neurons in primary visual cortex (Jones & Palmer, 1987). The primary goal of including the Gabor patch condition was to measure crowding over spatial locations with a commonly used simple stimulus to compare with letters. We predicted that the differences in crowding between upright letters and Gabor patches would be even stronger than the differences between upright and inverted letters. However, because letters and Gabor patches are very different stimuli, any observed differences in crowding distance may be due to differences in experience or to differences in low-level stimulus properties. Nonetheless, given the widespread use of Gabor patches in studies of crowding and other aspects of perception, it is worth testing whether and how crowding differs between Gabor patches and letters.

Data availability

Stimulus presentation scripts, data, and analysis scripts for Study 1 and Study 2 are available at https://osf.io/yw4mt/. This repository contains links to the preregistrations for both studies.

Method

Participants

The participants were undergraduate students enrolled in psychology courses at the University of California, Davis. All participants used for analysis were native and current readers of a Latin script (typically English or Spanish), native being defined as learned as their first written language or concurrently with their first written language. All participants reported having normal or corrected-to-normal vision. On the basis of a power analysis of data from a pilot experiment similar to Study 1a (N = 57) we preregistered a sample size of 60 usable participants for Studies 1a and 1b (see preregistration). Participants were considered usable if they met minimal standards for on-task performance as described below. All studies were approved by the Institutional Review Board at the University of California, Davis and all participants provided informed consent. In Study 1a, the average age of the 60 participants was 19.57 years (SD = 1.65). In Study 1b, the average age was 19.24 years (SD = 2.1). Additional demographic information for participants can be found in Table 1, which uses standard US-based categories for race and ethnicity (US Office of Management and Budget, 2016).

Materials and procedure

Stimulus presentation and data collection were controlled using PsychoPy (Pierce, 2007). Stimuli were presented on a HP ZR2440w LCD display with a gray background (26.8 cd/m^2) at a viewing distance of 100 cm. Viewing distance was measured for each participant and maintained by requiring participants to sit with their back touching the back of a chair which was placed such that the participant's eyes in this position were 100 cm from the center of the stimulus monitor. Experimenters carefully monitored participants to ensure that this viewing distance was roughly controlled.

Table 1
Gender, race, and ethnicity data for participants in Study 1a, Study 1b, Study 2.

Demographics of Participants Characteristic	Study 1a (n = 60)		Study 1b (n = 60)		Study 2 (n = 250)	
	n	%	*n*	%	*n*	%
Gender						
Female	47	78.3	45	75	56	22.4
Male	11	18.3	12	20	186	74.4
Other gender or not reported	2	3.3	3	5	8	3.2
Race						
Asian	27	45	27	45	119	47.6
White	23	38.3	20	33.3	78	31.2
Mixed race	4	6.6	5	8.3	15	6
Black	2	3.3	0	0	8	3.2
Native American	2	3.3	0	0	5	2
Other race or not reported	2	3.3	8	13.3	25	10
Ethnicity						
Not Hispanic or Latino	42	70	41	68.3	174	69.6
Hispanic or Latino	18	30	15	25	74	29.6
Not reported	0	0	4	6.6	2	0.8

The stimuli and procedure are illustrated in Fig. 2. Studies 1a and 1b were identical except that Study 1a compared upright letters (highly familiar stimuli) with inverted letters (less familiar stimuli), whereas Study 1b compared the same upright letters (highly familiar stimuli) with Gabor patches (unfamiliar stimuli). In each study, the two stimulus classes were factorially combined with 6 target locations on the visual field: 2, 4, and 6 degrees of visual angle to the left and right of fixation (on the horizontal meridian). The familiar upright letter stimuli were selected from a set eight white letters (F, G, J, L, N, S, R, Z) drawn in Arial font with a height of 1.0° and a luminance of 109 cd/m^2 (Fig. 2b). In Study 1a, the less familiar inverted letter stimuli were these same letters inverted (i.e., flipped vertically, not rotated 180 degrees – Fig. 2c). These specific letters were chosen to ensure that their upright and inverted forms would be different from each other. In Study 1b, the unfamiliar Gabor patch stimuli were selected from a set of eight Gabor patches (Fig. 2d) with a spatial standard deviation of ± 0.33° that varied in orientation (0°, 45°, 90°, or 135° from upright) and spatial frequency (4 or 6 cycles/°). In pilot testing, we found that these eight combinations of orientation and spatial frequency led to performance levels that were comparable to the eight inverted letters (which was not the case if we used eight different orientations with a single spatial frequency). Note that crowding is largely independent of the size of the target and flankers (Tripathy & Cavanagh, 2002), so we held stimulus size constant across eccentricities.

The task for each experiment was to identify the central stimulus in the array of three stimuli presented in each trial. For each of the twelve combinations of location and stimulus type, 50 trials were run, shuffled such that each condition was approximately evenly distributed across the 600 total trials. Each trial began with the subject fixating on a circular fixation point in the center of the screen, followed by a target appearing for 100 ms. The target was flanked horizontally by nontarget stimuli from the same set (e.g., nontarget inverted letters if the target was an inverted letter). The target-flanker distance was determined by a staircase procedure, described below. After 500 ms, a response array appeared, with the 8 possible stimuli from the stimulus set for that trial (e.g., the 8 inverted letters if it was an inverted trial, or the 8 Gabor patches if it was a Gabor trial), randomly distributed in a 4 × 2 grid on the screen. Subjects used a mouse to click on the stimulus that matched the target from that trial. Hence, the task was 8AFC for all experiments. After response, the fixation stimulus reappeared, and a new trial started after 1 s. For each experiment, participants completed a 24-trial practice version of the task before completing the actual experiment.

In each of the 12 conditions, the center-to-center distance between the target and each flanker was set according to the recommended "intensity" of a QUEST staircase (Watson & Pelli, 1983), targeting a recognition accuracy of 82 % correct for that condition. We chose 82 % as our threshold as this is a commonly used threshold for QUEST staircases. However, it should be noted that there is considerable variability in the choice of threshold used to measure crowding distances within the crowding literature, so care must be taken when comparing crowing distance estimates across studies (e.g. see Kurzawski et al., 2023). For all locations and stimulus conditions, the staircase had a prior of 0.5φ (i.e. spacing of 50 % of the target's eccentricity, φ) with an SD of.3φ. This prior was centered at 0.5φ because this represents the traditional estimate of the crowding distance (Bouma, 1970). Note that this measure of spacing abstracts away from the raw target-flanker distances and incorporates the linear eccentricity scaling that is fundamental to crowding ("Bouma's law"; Pelli & Tillman, 2008). For example, both a raw crowding threshold of 1° at an eccentricity of 4° and a raw crowding threshold of 0.5° at an eccentricity of 2° would each correspond to a scaled crowding threshold of 0.25φ. Thus, the null hypothesis for our eccentricity manipulation was that the scaled crowding threshold would be constant across eccentricities. The range of the QUEST staircase was limited to sample values between a maximum spacing of 1 and a minimum spacing of 0. The beta (shape) parameter for the staircase was fixed at 3.5, the gamma (lower asymptote) parameter was fixed at 0.125 (i.e., 1 divided by 8 possible answer choices), and the delta (upper asymptote) parameter was fixed at 0.02.

As a quality control measure, each session also contained 48 catch trials in which a target was presented without any flankers. These trials were distributed unpredictably across the session and included three trials for each of the 16 possible target stimuli (i.e. 3 for each letter and inverted letter/Gabor patch), with eight trials at each of the six locations. The catch trials were designed to be easy, so below-ceiling accuracy on these trials likely indicates either insufficient visual acuity or frequent lapses of attention. Participants were considered unusable if they missed more than two of the 48 catch trials. Participants were also considered unusable if more than two of the 12 staircases ended up at the maximum flanker distance (a crowding threshold of greater than 95 % spacing, which is implausible for individuals with typical vision). These exclusion criteria resulted in 7 unusable subjects in Study 1a, and 36 unusable subjects in Study 1b. Unusable participants were simply replaced so that we could reach the target N of 60 usable participants in each study. Individual staircases which were estimated to be near the maximum value were also excluded from analysis (0.83 % of staircases in Study 1a and 3.5 % of staircases in Study 1b). All of these data exclusion procedures were specified in the preregistration. Note that the increased number of excluded participants and staircases for Study 1b compared to Study 1a indicates that the inclusion of Gabor patches led to a more difficult task. Indeed, low performance in the Gabor patch conditions was the reason for the majority of excluded participants (33 out of 36) and excluded staircases (25 out of 26) in Study 1b. Hence, care should be taken when comparing the results of Study 1b with Study 1a, as the exclusion criteria may have in effect caused a more stringent sample selection for Study 1b. An analysis of the isolated catch trials, which shows the uncrowded recognition accuracy for stimuli used in Study 1, is available in the supplemental materials.

Data analysis

The dependent variable was the *crowding threshold*, defined as the target-flanker spacing (expressed as proportion of the target's eccentricity) required to achieve 82 % accuracy. This value was computed as the mean of the posterior distribution for the 50-trial QUEST staircase for each condition. Hence, each participant contributed 12 values to the analysis – one crowding threshold for each combination of the two stimulus conditions (upright letters and inverted letters/Gabor patches), three eccentricities (2°, 4°, and 6°), and two hemifields (left and right).

These data were modeled separately for each experiment using linear mixed effect regression as implemented in the lme4 R package (Bates, et al., 2014). The structure of the model included fixed effects for Eccentricity, Hemifield, and Stimulus, and all possible interactions between these three variables. Hemifield and Stimulus were effect coded (i.e., −0.5 and 0.5 for left vs. right; −0.5 and 0.5 for upright letters vs. inverted letters or Gabor patches). Eccentricity was coded as a continuous variable centered at the intermediate eccentricity (i.e., 2°, 4°, and 6° were coded as −0.5, 0, and 0.5, respectively). The random effect structure included random intercepts for Subject, as well as by-subject random slopes for the effect of each of the fixed effects. In other words, the model did not assume that the slopes were identical across participants. As specified in the preregistration, the random slopes for the interactions between the fixed effects were dropped from all models to allow them to converge. Significance for each effect was tested with a Type 3 Wald chi-square test using Satterthwaite's method through the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017). These tests were preregistered and used uncorrected *p* values with an alpha of.05. Beta weights and standard errors are also reported from the models; their units being spacing as a ratio of the target's eccentricity (φ). Note that because the models were effect-coded, the model beta coefficients for the effects are directly interpretable as the predicted difference in scaled crowding distance due to that effect over the average of the other variables. For Hemifield, this would be the average change in φ for locations in the right hemifield compared to the left; for Stimulus, this would be the average change in φ for inverted letters or Gabor patches compared to upright letters; Eccentricity, this would be the average change in φ for the 6° eccentricity compared to the 2° eccentricity location.

To aid the interpretation of interaction effects, follow-up analyses were conducted by fitting the same model separately for each stimulus type without the fixed or random effects of Stimulus. Additionally, a follow up model was fit to compare the inverted letters from Study 1a with the Gabor patches in Study 1b. This model had the same structure as the main models, but without random slopes for Stimulus (because Stimulus was a between-subjects variable for this analysis). Full model outputs for the main and follow up models are available in the supplementary materials. The supplementary materials also contain alternative versions of the models from Study 1, which use dummy-coded factors, test interactions from the unified model, and include estimated marginal means for each condition. These models have identical patterns of significance to the models reported in the manuscript, but may be useful to researchers extract the crowding distances for each condition in the model.

When an error is made during a crowding task, participants often report the identity of a flanker. This is known as a substitution error or source confusion. The present study was not designed to analyze substitution errors, but for the curious reader, an exploratory analysis of the substitution errors from Study 1 is available in the supplementary materials.

Results and discussion

Study 1a: Upright versus inverted letter crowding

Fig. 3a shows the mean eccentricity-scaled crowding thresholds for each cell of Study 1a. For upright letters, mean crowding distances ranged from about 0.27φ at the 2° eccentricities to about 0.43φ at the 6° left location. Inverted letters had a smaller range in crowding distances, with a mean crowding distance of about 0.43φ at the 2° left location and a distance of about 0.56φ at the 6° left location. The intercept of the model for Study 1a was 0.417φ, indicating the grand mean of the crowding distances across all conditions. The crowding threshold was on average 0.154φ lower for upright letters than for inverted letters across the locations tested ($\beta = 0.154$, $SE = 0.009$, $p < 0.001$). Thus, although the upright and inverted letters were essentially identical in terms of their visual properties, the upright letters could be identified at 82 % accuracy with closer spacing than the inverted letters. This provides evidence for an effect of experience on the distance at which flanking objects disrupt the identification of a target object, which is perhaps the most important characteristic of crowding.

Crowding thresholds were lower when the target was closer to fixation than when the target was farther away ($\beta = 0.113$, $SE = 0.011$, $p < 0.001$), with an increase of crowding distance of 0.113φ from the 2° eccentricity locations compared to the 6° eccentricity locations. Note that this is despite the fact that these distances were already were scaled to be proportional to the eccentricity. The eccentricity effect was steeper by 0.058φ for the upright letters compared to the inverted letters, corresponding to a significant Stimulus x Eccentricity interaction ($\beta = -0.058$, $SE = 0.016$, $p < 0.001$). Thus, the eccentricity function of crowding—perhaps the second most important characteristic of crowding—also differed between highly familiar upright letters and less familiar inverted letters.

We decomposed the Stimulus × Eccentricity interaction by testing

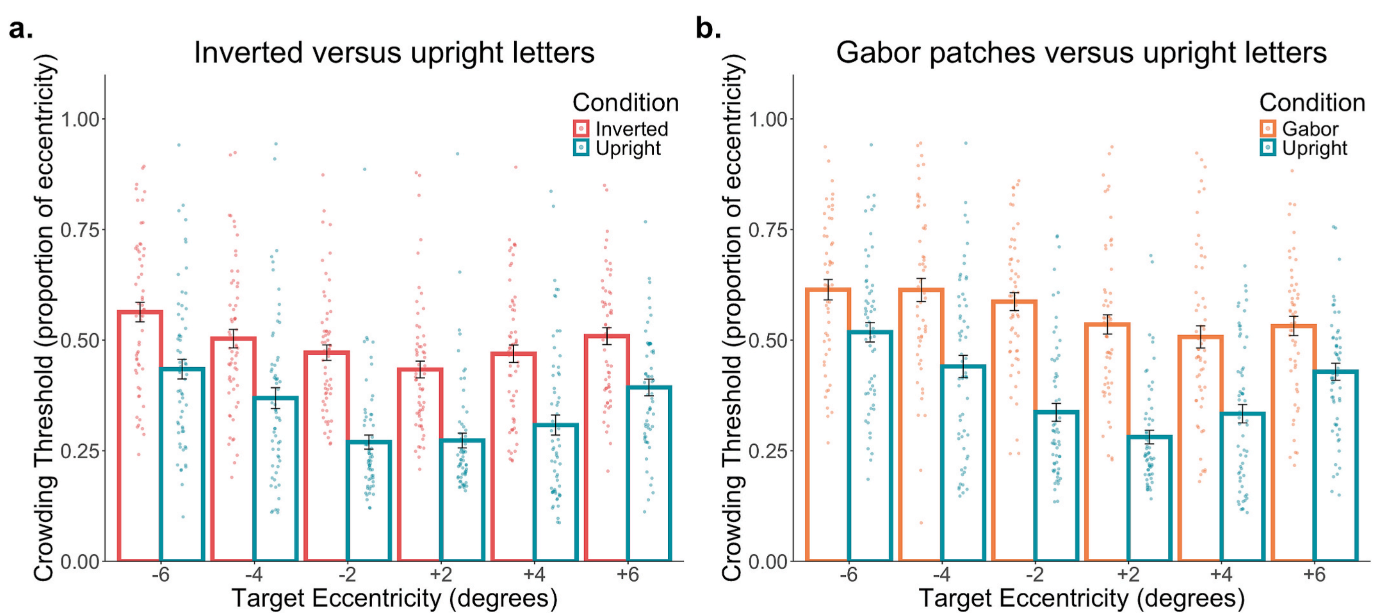

Fig. 3. Mean eccentricity-scaled thresholds for each location and familiarity condition, comparing upright and inverted letters in Study 1a (**a**) or upright letters and Gabor patches in Study 1b (**b**). Error bars show ± 1 SEM. Dots show single-participant thresholds. Note that a perfectly linear relationship between eccentricity and crowding threshold would lead to a flat function over eccentricities (e.g., all values ~ 0.5φ).

the eccentricity effect separately for upright letters and for inverted letters. The Eccentricity effect was significant for both upright letters ($\beta = 0.142$, SE $= 0.014$, $p < 0.001$) and inverted letters ($\beta = 0.083$, SE $= 0.011$, $p < 0.001$) when they were analyzed separately. Thus, even though the eccentricity effect was greater for upright than inverted letters, the eccentricity-scaled thresholds increased with eccentricity for both stimulus types. This is a violation of Bouma's law, which specifies that the eccentricity-scaled threshold should be constant (i.e., that the absolute threshold should be a constant proportion of the eccentricity).

As illustrated in Fig. 3a, crowding thresholds were 0.036φ smaller for targets to the right of fixation than to the left ($\beta = -0.036$, SE $= 0.012$, $p = 0.005$). This was predicted because experienced readers of Latin scripts regularly sample letter information to the right of fixation prior to each saccade. However, the Stimulus × Hemifield interaction was near zero and not statistically significant ($\beta = -0.008$, SE $= 0.013$, $p = 0.564$), suggesting a general advantage for the right visual field rather than a specific rightward bias for upright letters.

Study 1b: Upright letter versus Gabor patch crowding

Fig. 3b shows the mean eccentricity-scaled crowding thresholds for each cell of Study 1b. For upright letters, mean crowding distances ranged from about 0.28φ at the 2° right location to about 0.52φ at the 6° left location. Gabor patches had a smaller range, with the lowest mean crowding distance of about 0.51φ at the 4° right location and a distance of about 0.61φ at the 6° left location. The overall intercept of the model for Study 1b was 0.481φ. Upright letters had overall crowding thresholds 0.183φ lower than Gabor patches ($\beta = 0.183$, SE $= 0.016$, $p < 0.001$). This provides additional evidence that the visual system can tolerate closer spacing for familiar upright letters than for other stimuli. At this point, however, we cannot determine whether the difference between letters and Gabor patches reflects an effect of experience or an effect of other differences in stimulus properties. We also compared the data from the inverted letters from Study 1a with the data from the Gabor patches from Study 1b. We found that thresholds were 0.079φ lower for the inverted letters than for the Gabor patches ($\beta = 0.079$, SE $= 0.022$, $p < 0.001$).

As in Study 1a, eccentricity-scaled thresholds for upright letters in Study 1b decreased substantially as target eccentricity decreased, violating Bouma's law. However, Gabor patches followed Bouma's law, exhibiting little or no effect of eccentricity on the eccentricity-scaled thresholds. This pattern led to a significant Stimulus x Eccentricity interaction ($\beta = -0.149$, SE $= 0.023$, $p < 0.001$), with upright letters showing 0.149φ more eccentricity scaling between the closest and furthest eccentricities compared to Gabor patches. When analyzed separately, the effect of Eccentricity was significant for letters ($\beta = 0.165$, SE $= 0.016$, $p < 0.001$) but not for Gabor patches ($\beta = 0.016$, SE $= 0.018$, $p = 0.386$).

Additionally, an across-experiment follow-up analysis indicated that the effect of eccentricity was significantly greater for inverted letters than for Gabor patches (Stimulus × Eccentricity interaction: $\beta = -0.067$, SE $= 0.022$, $p = 0.002$). In other words, Bouma's Law was confirmed for Gabor patches, but it was violated for inverted letters and violated even more strongly for upright letters. These results mirror the pattern of effects observed for the overall thresholds, with the largest eccentricity effect for the upright letters, weaker eccentricity scaling for the inverted letters, and no eccentricity scaling beyond Bouma's law for the Gabor patches. The violation of Bouma's law for letters but not for Gabor patches suggests a fundamental difference between Gabor patches and letter-like stimuli in the nature or effects of crowding.

Finally, crowding thresholds were 0.088φ lower for targets in the right visual field than for targets on the left ($\beta = -0.088$, SE $= 0.016$, $p < 0.001$). However, the interaction between Stimulus and Hemifield was near zero and not significant ($\beta = -0.007$, SE $= 0.019$, $p = 0.691$), consistent with a general advantage for the right visual field.

Study 2: Individual differences in visual crowding

In Study 1, we found several effects consistent with the hypothesis that reading experience leads to changes in the effects of visual crowding. Most importantly, crowding thresholds were lower for upright letters than for inverted letters. In addition, whereas Gabor patches followed Bouma's law, with no change in eccentricity-scaled thresholds across eccentricities, these thresholds were reduced near fixation for letters, and more so for upright than inverted letters. This eccentricity effect fits with the fact that reading is primarily performed at or near fixation. Finally, crowding was reduced in the right visual field (for all stimulus types), matching the fact that information from the right visual field plays an important role in guiding saccades during the reading of languages that use the Latin script. Although the reduced crowding effects in the right visual field for all stimulus types may not necessarily be attributed to reading experience, the differences between upright and inverted letters in overall thresholds and eccentricity effects were almost certainly a result of reading experience given that inverted letters are almost identical to upright letters in every property other than experience. In Study 2, we aimed to provide additional evidence that crowding for upright letters is related to reading experience, using a large-N individual-differences approach. Specifically, we asked whether intersubject variability in reading experience (as assessed by two proxy measures) is associated with the putative reading-related properties of crowding that we observed for upright letters in Study 1 (reduced thresholds, a drop in eccentricity-scaled thresholds near fixation, and a right hemifield advantage). Our primary proxy measure—the Andrews spelling test—quantifies the ability to recognize misspelled words and is thought to reflect the quality or precision of word representations, which are developed through reading practice (Andrews, Veldre, & Clarke, 2020). Our secondary proxy—the author recognition task—quantifies the number of author names that an individual recognizes, which is an indirect measure of lifetime reading experience (Acheson, Wells, & MacDonald, 2008). We predicted that individuals who scored higher on these tests would have decreased crowding thresholds, especially for locations closer to fixation and locations in the right visual field.

Although we found an equivalent right hemisphere advantage in Study 1 for upright letters, inverted letters, and Gabor patches, this effect might still be a consequence of reading experience. For example, massive experience with left-to-right reading might cause general changes in attention that impact all stimuli and not just letters. If so, then we would expect that the magnitude of the right hemisphere advantage would be correlated with our proxies for reading experience in Study 2.

Additionally, in Study 2 we measured visual working memory capacity with a change localization task that provides a sensitive and reliable measure of general visual and cognitive ability (Johnson et al., 2013). This allowed us to assess the possibility that the relationship between our reading experience measures and crowding reflects some broader aspect of visual cognition rather than being specific to reading.

Method

Participants

The original preregistered sampling plan had the same inclusion criteria as in Study 1, including all individuals who were native and current readers of any language using a Latin script. On the basis of a power analysis using data from 115 participants in pilot study, we preregistered a sample size of at least usable 200 participants, with data collection continuing until the end of the spring 2022 term (see preregistration). This led to a sample of 267 usable subjects. However, we later realized that our measures of reading experience—which are based on English-language spelling ability and the recognition of English-language authors—were designed for native readers of English and have not been validated for native readers of other languages. We therefore limited inclusion to native and current readers of English (N = 250) and excluded 17 individuals who were native and current readers of other languages that use a Latin script (primarily Spanish). This difference did not impact the statistical significance of any of the effects, but it is nonetheless the more appropriate sampling method, so the analyses provided here were limited to the 250 native readers of English. Analyses of the full sample of 267 participants are provided in the supplementary materials (see Table S18). All participants reported having normal color vision, which was required due to our visual working memory task which was a color change localization task (described below). The average age of the 250 participants was 20.04 years (SD = 2.27). Additional demographic information can be found in Table 1.

Materials and procedure

All participants performed four tasks in a fixed order: a spelling task, an author recognition task, a crowding task, and a visual working memory task. The spelling and author recognition tasks were run first to ensure that they would not be contaminated by an individual's perceived performance on the crowding task. The working memory task was run last because it was considered secondary. In addition, it was important to keep the order constant across participants to avoid adding an uncontrolled source of variance among individuals.

Crowding task

The crowding task for Study 2 was identical to that for Study 1 except as follows. Only familiar upright letters were used. The set of letters used was the eight consonants on the middle row of a standard keyboard (S, D, F, G, H, J, K, L). These letters were used because in Study 2, participants responded by pressing the corresponding key on the keyboard rather than reporting the perceived target using a mouse. These eight letters allowed participants to easily keep their fingers on the response keys. This approach yielded faster behavioral responses and allowed us to replicate the pattern of letter crowding results from Study 1 with a different set of letters. Because of the faster behavioral responses and the reduced number of conditions in the crowding task for Study 2, we observed in our pilot study that this version of the task could be completed much faster. This allowed us to raise the number of trials for each condition from 50 to 70, to further increase the precision of the single-participant crowding thresholds. Consequently, the total number of trials for the crowding task in Study 2 was 420. Participants completed an 18-trial practice version of the task before completing the actual experiment. A total of 42 catch trials were presented, including seven trials at each of the six locations. As specified in the preregistration, participants were considered unusable if they missed more than two of these 42 catch trials or if more than one of the six staircases ended up at the maximal flanker distance. These exclusion criteria resulted in 39 subjects being categorized as unusable and therefore being replaced. Individual staircases which were excluded for being near the maximum value amounted to 0.6 % of staircases in Study 2.

Visual working memory task

Visual working memory capacity was measured using a color change localization task with a set size of 6 items, presented on the same video monitor used for the crowding task. Change localization performance is substantially more reliable than performance on the more common change detection task (Kyllingsbaek & Bundesen, 2009; Zhao, Vogel, & Awh, 2022), is strongly correlated with change detection performance (Zhao et al., 2022), exhibits excellent test–retest reliability (Johnson et al., 2013), and is strongly correlated with IQ and other measures of overall cognitive ability (Johnson et al., 2013). For each trial in this task, six colored disks (each with a radius of half a degree of visual angle) appeared for 200 ms, equally spaced around a notional circle with an eccentricity of 3° from a central fixation cross. After a retention period of 900 ms, the disks reappeared, with a disk at one randomly selected location having changed color. The participants then used a computer mouse to click on the changed disk. The colors of the disks were randomly drawn from sets of approximately isoluminant colors, such that the colors in a given array were roughly equally spaced from across the color wheel. The color of the changed disk was 120 degrees on the color wheel away from the original color. Participants completed 96 trials after experiencing 12 practice trials. The dependent measure for this task was the number of trials with a correct response, which is perfectly correlated with the traditional K measure of storage capacity for this task (Johnson et al., 2013). For our sample of participants, the median raw score for the visual working memory task was 48 (Mean = 48.66, SD = 10.88). The split-half reliability of this task in our sample of participants was high (r = .80).

Spelling task

Our primary proxy for reading experience was the Andrews spelling task, which measures a participant's ability to identify misspelled words and has been shown to have high test–retest reliability (r = .93; Andrews & Hersch, 2010). Spelling ability is thought to represent the quality or precision of a person's lexical representations, which are learned through experience (for a review see Andrews et al., 2020). We used an updated version (Andrews et al., 2020). Participants were given a sheet of paper with 88 words arranged in a grid. Half of the words were misspelled, and the participant was instructed to circle these words. The score was the number of circled misspelled words plus the number of uncircled correctly spelled words. For our sample of participants, the median raw score for the spelling task was 76 (Mean = 74.48, SD = 7.86).

Author recognition task

Our secondary proxy for reading experience was the *author recognition task*. This task is designed to quickly measure print exposure, based on the assumption that people who have read more will be able to recognize more authors (Stanovich & West, 1989). We used a version from 2008 (Acheson, et al., 2008), in which 130 real and fake author names were intermixed in a grid on a sheet of paper. Participants circled all the names they recognized as being authors. Participants received one point for each correct author circled and were penalized one point for each non-author circled, meaning the maximum possible score was 65. Meta-analysis of studies using this task show it has good reliability (Cronbach's α = .75–.89; see Mol & Bus, 2011). However, note that this task is not a very good measure for people who come from non-Western backgrounds (McCarron & Kuperman, 2021), like many of the students at the University of California, Davis. Consequently, we did not expect it to be as good of an indicator of reading experience as the Andrews spelling task. Further, it has also been found that this task may not be an ideal measure for current college undergraduates, as the average score among students is quite low (Moore & Gordon, 2015). For our sample of participants, the median raw score for the author recognition task was 8 (Mean = 9.28, SD = 5.81).

Data analysis

The analysis strategy was patterned after that used in Study 1. Crowding thresholds at each of the 6 locations were modeled with fixed effects for Hemifield, Eccentricity, and the interaction between the two. Additionally, and interacting with each of these effects, there were fixed effects for the three individual difference measures –visual working memory, spelling, and author recognition. Hence the highest order interactions were 3-way interactions between Hemifield, Eccentricity, and each of the 3 individual difference measures (as specified in the preregistration). Each of the individual difference measures was normalized across the sample of subjects, putting all three measures onto the same scale of z scores. In the present sample, the correlation between spelling and author recognition was 0.328, the correlation between visual working memory and spelling was 0.109, and the correlation between visual working memory and author recognition was −0.016. The random effect structure included random intercepts for subjects and by-subject random slopes for Hemifield and Eccentricity. Significance testing was identical to the models in Study 1. The full model output is available in supplementary materials (see Table S17).

As with Study 1, an analysis of the substitution errors made in Study 2 and how their likelihood is predicted by scores on the individual difference measures is available in the supplementary materials.

Results and discussion

As shown in Fig. 4, mean crowding distances for letters in Study 2 ranged from about 0.28φ at the 2° right location to about 0.59φ at the 6° left location. The intercept of the model was 0.434φ, indicating the grand mean of crowding distance across all locations, for an average score on all individual difference measures. The eccentricity-scaled crowding thresholds were significantly smaller when the targets were closer to fixation (β = 0.211, SE = 0.007, $p < 0.001$), with crowding distances for the 2° locations being 0.211φ lower than for the 6° locations. Crowding distances were also 0.099φ lower for targets to the right of fixation than for targets to the left (β = −0.099, SE = 0.008, $p < 0.001$). This replicates the pattern observed for upright letters in Study 1.

To visualize the correlations between crowding thresholds and the individual difference measures, we plotted the estimated slope of the relationship between eccentricity-scaled crowding thresholds and the spelling task (Fig. 5a), the author recognition task (Fig. 5b), and the visual working memory task (Fig. 5c), averaged over the 6 locations. A negative slope indicates less crowding (a lower threshold) for people with higher scores on these variables. To visualize the effect of each individual difference measure over the 6 locations, we plotted the model-predicted change in crowding threshold for an increase of one standard deviation on the spelling task (Fig. 5d), the author recognition task (Fig. 5e), and the visual working memory task (Fig. 5f), for each of the 6 locations. In Fig. 5g-i, we also plotted these same model-predicted changes for models fit on each individual difference measure separately. These show the effects of each variable without controlling for the effects of the other individual difference variables, to aid in the interpretation of the individual effects and give a sense of the degree of the collinearity between variables. The results of these models had the same pattern of significance as the main model.

Fig. 5a shows that higher scores on the spelling test were associated with significantly lower crowding thresholds across the six locations (β = −0.022, SE = 0.007, p = 0.003), with an increase of one SD in spelling performance predicting a decrease of 0.022φ in crowding distance. Fig. 5d shows that this effect was especially prominent when the target was presented to the right of fixation, leading to a Spelling × Hemifield interaction (β = −0.018, SE = 0.008, p = 0.027). Thus, we found that greater spelling ability—which is a proxy for reading experience—significantly predicted reduced overall crowding thresholds and a greater right hemifield advantage. These results provide further evidence for the hypothesis that greater reading experience is associated with reduced crowding thresholds for letters, especially in areas of the visual field where readers have the most experience recognizing crowded letters. Note that performance on the spelling test had a numerically larger impact for stimuli closer to fixation, but the interaction with Eccentricity was not significant (β = 0.010, SE = 0.007, p = 0.164). It is possible the association between Spelling and Eccentricity would have reached significance if we had included a broader range of eccentricities or if we had tested an even larger sample of subjects.

As shown in Fig. 5b, performance on the author recognition task did not predict overall crowding thresholds (β = 0.001, SE = 0.007, p = 0.904). However, there was a significant three-way interaction of Author Recognition × Hemifield × Eccentricity (β = 0.031, SE = 0.012, p = 0.013). As shown in Fig. 5e, this interaction appeared to reflect increased scores on author recognition being associated with lower crowding thresholds for targets specifically in the right hemifield and at the nearest eccentricity. One reason for the limited effect of author recognition may be that this measure was positively correlated with spelling ability (*r* = .33 in our sample), and variance in crowding was attributed to spelling ability. Consistent with this, when spelling was removed from the model, the effect of author recognition increased markedly (but was still focused on the locations immediately to the right of fixation – see Fig. 5h).

To examine whether crowding thresholds are related to reading experience per se or whether associations would be seen for any reliable measure of cognitive ability, we assessed whether individual differences in visual working memory performance showed the same pattern of association with crowding thresholds. As illustrated in Fig. 5c, greater visual working memory scores were associated with significantly lower overall crowding thresholds (β = −0.034, SE = 0.007, $p < 0.001$), with an increase of one SD in visual working memory performance predicting a decrease of 0.034φ in crowding distance. However, as shown in Fig. 5f,

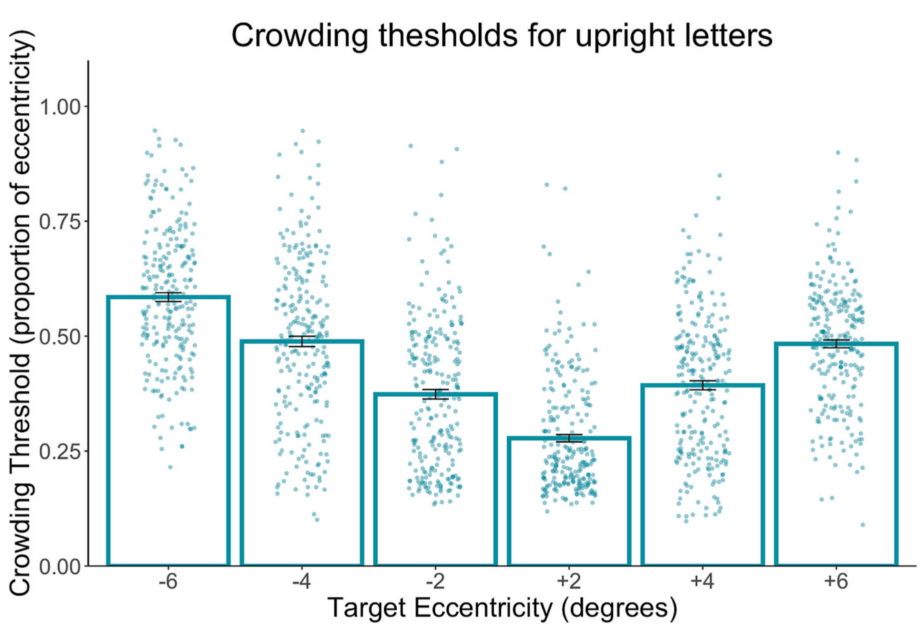

Fig. 4. Mean eccentricity-scaled thresholds for upright letters at each location in Study 2. Error bars show ± 1 SEM. Dots show single-participant thresholds. Note that a perfectly linear relationship between eccentricity and crowding threshold would lead to a flat function over eccentricities (e.g. all values ~ 0.5φ).

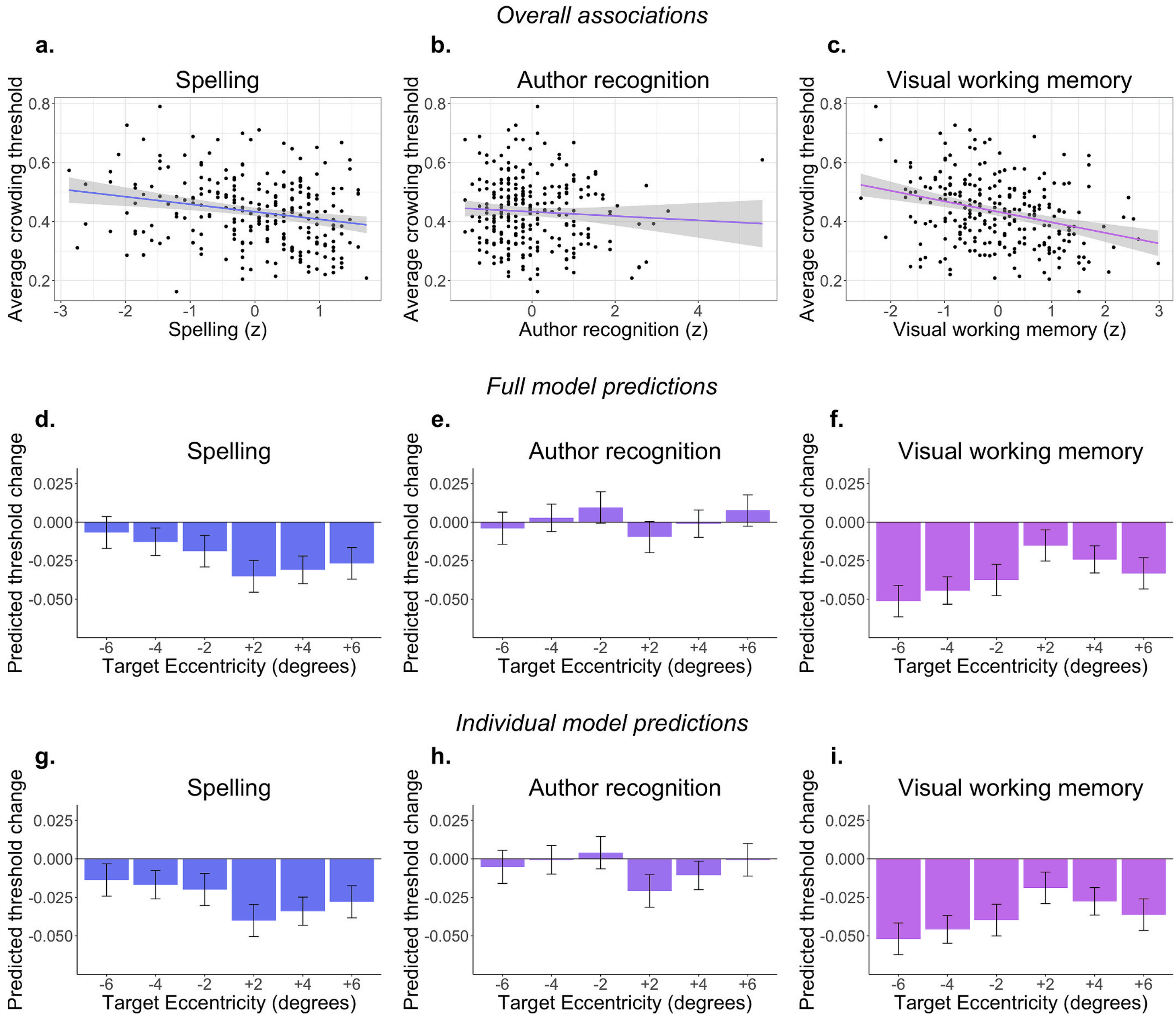

Fig. 5. a-c. Relationship between average crowding threshold (over the 6 locations) and each of the individual difference measures (Spelling, Author recognition, and Visual working memory). Dots are individual subjects, and the lines are simple linear regression fits. The shaded region shows the standard error of estimate for the regression line at each point. **d-f.** Model-predicted change in crowding threshold for an increase of 1 standard deviation on each of the individual difference measures, per location, in the full model (i.e., the effect of one variable after controlling for the other variables). **g-i.** Model-predicted change in crowding threshold for an increase of 1 standard deviation on models fit with each of the individual difference measures separately. Error bars show ± 1 SEM.

the spatial pattern was the opposite of the effects of spelling ability. Specifically, the visual working memory effects were larger to the left of fixation ($\beta = 0.020$, SE = 0.008, $p = 0.010$) and larger at greater eccentricities ($\beta = -0.016$, SE = 0.007, $p = 0.024$). Thus, although both visual working memory performance and spelling performance were associated with decreased crowding thresholds, each explained unique variance and had a different spatial pattern, suggesting different underlying mechanisms.

General discussion

Letters are tightly packed within words, and the reduction in perceptibility produced by this crowding is a major limitation on reading (Pelli et al., 2007). The present results demonstrate that the human visual system minimizes the deleterious effects of crowding in three ways that match the specific demands of reading. First, the visual system can tolerate closer spacing for upright letters than for inverted letters, which matches the fact that almost all reading is done with upright letters. Second, the effects of crowding on upright letters are increasingly minimized for stimuli closer to fixation, which matches the fact that reading puts high demands on the identification of tightly spaced but highly similar stimuli near the fovea. Third, the visual system can tolerate closer spacing of stimuli in the right visual field than in the left visual field, which matches the fact that readers of Latin scripts make extensive use of parafoveal information in the right visual field (Schotter et al., 2012; Veldre & Andrews, 2016; Rayner et al., 2011). In addition, people who scored higher on the Andrews spelling test—which appears to reflect experience-driven changes in the precision of visual word representations (Andrews et al., 2020)—could tolerate closer spacing for upright letters, especially in the right visual field.

The differences in crowding effects between upright and inverted letters were almost certainly caused by our participants' massive experience with reading, and we would expect that most literate adults in industrialized societies would exhibit similar effects. Although it is not practical to experimentally manipulate experience at this scale in humans, the observed differences in crowding between upright and inverted letters would be difficult to explain by anything other than differences in experience. That is, given that upright and inverted letters are almost identical in terms of their physical properties, it is implausible that people without experience reading the Latin script would show different crowding patterns for upright and inverted letters from this script. Thus, the results of Study 1 provide strong evidence that reading experience leads to the ability to recognize upright letters well when they are closely surrounded by other nearby letters. That is, reading experience minimizes the deleterious effects of crowding on letter perception.

Interestingly, Studies 1a and 1b indicated that the eccentricity profile of crowding varies systematically across experience and/or stimulus type. We found that Gabor patches obeyed Bouma's Law (Pelli & Tillman, 2008), which states that the critical distance for crowding remains constant when expressed as a proportion of target eccentricity. This is consistent with the few previous studies which have directly assessed the crowding distance at multiple eccentricities for Gabor patches (e.g. Levi & Carney, 2009). However, we found that Bouma's Law was clearly violated by letters, which yielded lower-than-expected thresholds near fixation. This violation was stronger for upright letters than for inverted letters, indicating that the deviation from Bouma's law was at least partly a result of experience. This finding does conflict with the commonly held idea that the crowding distance for letters is a simple linear function of the eccentricity of the target (Bouma, 1970; Pelli & Tillman, 2008). Note that while the current study found that this was not the case for horizontal letter crowding along the horizontal meridian in the parafovea, our experiments did not test eccentricities beyond 6°. It may be likely that Bouma's law holds for letters beyond the parafovea and in other spatial locations that are not as critical for reading.

Further research would be needed to determine whether the differences in overall crowding thresholds between letters and Gabor patches are a result of experience or a result of the physical properties of the stimuli. For instance, crowding distances for Gabor patches may be greater compared to letters because they are visually simpler, because they have higher target-distractor similarity, or because their features are encoded along continuous dimensions (e.g. orientation). These differences may to lead crowded arrays of Gabor patches being more readily pooled or averaged within the receptive fields of neurons in early visual cortex, which is indeed the primary mechanism in many models of visual crowding (e.g. Levi, 2008, Pelli & Tillman, 2008). This may also contribute to the more typical pattern of eccentricity scaling observed for Gabor patch crowding compared to that of upright letters. Additionally, as discussed in the Methods section, more participants were excluded for poor performance on Gabor patch catch trials than letter trials, suggesting that the task for Study 1b was more challenging than for Study 1a. This suggests that the Gabor patch stimuli may have been inherently more difficult to encode or remember than the letter stimuli, perhaps due in part to letters having verbal labels associated with them (that are learned through experience).

However, the low-level similarity structure was highly controlled between the upright and inverted letters used in Study 1a. Hence, the difference in both crowding distance and eccentricity scaling we observed between upright and inverted letters is strong evidence that crowding is affected by experience. These results fit most naturally with models of crowding that allow for crowding effects to arise from interactions at multiple levels in the visual hierarchy (e.g., Manassi & Whitney, 2018). One such model is the hierarchical sparse selection model (Chaney, Fischer & Whitney, 2014) in which crowding for objects is also a function of the receptive fields of the populations of neurons that represent the object itself. This model predicts differences in crowding across stimulus types, and accounts for other related results such as the finding that upright faces are crowded more by other upright faces than inverted ones (Louie, Bressler & Whitney, 2007). Under this framework, the constraints of reading may spatially bias the tuning of letter representations such that their receptive fields are smaller, particularly closer to the fovea, causing letters to experience reduced crowding. The finding that inverted letters also deviate from Bouma's law may indicate that some of this tuning may be occurring at the level of letter features as well.

A very similar idea has also been put forward as the modified receptive field hypothesis (Chanceaux & Grainger, 2012), in which the receptive field properties of neurons involved in letter recognition are modified by reading to reduce interference from neighboring letters and optimize the parallel recognition of letters. It is important to note that for both previous results of increased recognition accuracy for crowded letters (e.g. Tydgat & Grainger, 2009) and the current finding of reduced crowding thresholds for letters, it is unknown whether the fundamental mechanism of crowding is altered for letters, or there is some other process which allows for upright letters to be recognized under crowded conditions. For example, if experience with upright letters simply makes them more discriminable in the presence of noise, they may be more easily perceived than inverted letter in the presence of noise from flanking letters, resulting in reduced crowding thresholds.

Future research will be needed to distinguish the various possible mechanisms for these effects. Indeed, they are not mutually exclusive, and the mechanism or mechanisms of visual crowding are still not fully understood (e.g. Manassi & Whitney, 2018). Whatever the underlying mechanism, however, the present results show that crowding is *functionally* reduced for upright letters and that Bouma's law—which we found to be true for Gabor patches—is far from an accurate description of crowding for upright letters. This finding informs models of visual crowding, and suggests that differences in stimulus classes, both in terms of their physical characteristics as well as the observer's experience with the stimulus, should be taken into account. Further, these findings help to characterize the extent of interference between letters as a function of spacing over eccentricities and hemifield, which is informative for models of reading – particularly for models which incorporate positional letter representations that depend on their eccentricity and hemifield (e.g. Snell, 2024).

Although conclusions about causation are more difficult for the correlational approach used in Study 2, the observed association between spelling performance and crowding thresholds, particularly in the right visual field, is at least consistent with the proposal that reading experience alters letter crowding. Performance on the author recognition test was not significantly related to overall crowding thresholds, although it was significantly associated with lower crowding thresholds in the near-right parafovea, a key location for readers of Latin scripts. Hence, this effect is also consistent with a role of reading experience on sensitivity to crowding by nearby letters. These effects were not driven by greater overall cognitive ability in people who scored higher on the two reading experience scales, because a very different pattern of associations was observed for an independent measure of working memory capacity. While we suggest there is good reason to believe that reading experience would cause changes in letter crowding effects, there is also the possibility that the causal direction is flipped or mediated by a third variable. That is, having lower crowding thresholds for letters (perhaps due to some factor unrelated to reading) may lead people to spend more time reading, leading to the correlations we observed. In either case, these findings may be relevant to researchers who are developing procedures for improving reading ability. That is, they indicate that experienced readers have an advantage in the perception of crowded letters, which presumably aids in their reading ability. Conversely, inexperienced readers may face increased perceptual challenges because they are more susceptible to crowding. This suggests that care should be taken to ensure that inexperienced readers are given reading materials in which the letters are widely spaced.

We would like to stress that although our proxies for reading experience are widely used, they are far from perfect measures of actual reading experience. Reading is a complex human activity that involves the integration of many visual and cognitive skills. Spelling ability is thought to index the quality or precision of visual word representations which are learned through experience (Andrews et al., 2020). This is an important skill for reading, but only one aspect of many. Author recognition is thought to index time spent reading more generally (Acheson et al., 2008). However, this measure likely does not fully capture all types of reading, for instance time spent reading in many online contexts or as subtitles in videos. An interesting avenue for future research would be to obtain more measures of reading skill and experience and assess how they are associated with visual crowding. Another interesting test of the effects of reading on crowding would be to examine letter crowding for readers of languages with different scanning patterns than English. On the basis of the current findings, we would predict that experience reading right-to-left scripts (e.g., Arabic or Hebrew) might create a bias toward the left visual field instead of the right visual field. Similarly, experience reading more complex logographic scripts (e.g., Chinese)—where there is more information concentrated within a fixation—may change the nature of the eccentricity scaling for letter crowding.

We found smaller crowding thresholds in the right visual field than in the left visual field in all three experiments. This effect was similar in magnitude for upright letters, inverted letters, and Gabor patches, so it is not clear whether this effect is related to reading experience. It could be an experience-independent bias that is related to intrinsic differences between the left and right hemispheres (e.g., Michael & Ojeda, 2005), although interestingly, crowded letters in the left visual field benefit more from attentional cues than crowded letters in the right visual field (Ramamurthy et al., 2021). Another factor is that words are ultimately processed in the left hemisphere, so letters in the right hemifield may receive privileged processing (Yeatman & White, 2021). However, this would not explain the fact that a right hemifield advantage was also observed for Gabor patches.

Alternatively, experience with reading left-to-right scripts could produce a generalized attentional bias for all stimulus types. This alternative is consistent with the finding from Study 2 that the right hemifield advantage was greater in people with higher scores on both the spelling test and the author recognition test. Previous research shows that readers have asymmetric perceptual span sizes biased towards their reading direction, for instance to the right in English but to the left in Hebrew (Pollatsek et al., 1981) or Urdu (Paterson et al., 2014). Further, in English readers, increased spelling ability and reading comprehension are associated with larger perceptual spans in the right visual field (Veldre & Andrews, 2014). One possibility is that the development of visual span and parafoveal processing during reading in children (Kwon, Legge, & Dubbels, 2007) occurs in tandem with the development of mechanisms to reduce crowding. One mechanism for reduced crowding in the right visual field may be alterations in inward-outward asymmetry, a characteristic of radial crowding where an outer flanker impairs the recognition of target more than an inner flanker (Petrov, Popple & McKee, 2007). Interestingly, it has been found that this asymmetry is reduced in the right visual field compared to the left visual field, potentially contributing to reduced overall crowding in the right visual field (Chakravarthi et al., 2021). One explanation for this difference may be related to the model of crowding put forth by Nandy and Tjan (2012), where the tuning of lateral interaction zones outside of the fovea is learned though spatial attention but biased by eye movements. Given that English reading puts particular constraints on spatial attention and eye movements that are focused in the right visual field, it is possible that this learning process is a source of the general bias of lower crowding distances in the right visual field.

The large sample size of 250 observers in Study 2—to our knowledge, the largest sample in any crowding study published to date—made it possible to demonstrate that crowding varies considerably across individuals and is associated with high-level factors such as spelling ability and working memory capacity. The relationship was relatively strong for working memory capacity, with the model suggesting an overall decrease in crowding threshold of about.034φ for each increase of one standard deviation of working memory capacity. Clearly, observers may leverage visual working memory to retain target identity before responding. However the fact that higher visual working memory capacity predicted lower crowding distances suggests that visual working memory may be more directly involved in disambiguating between target and flanker stimuli from crowded visual representations. This effect is consistent with electrophysiological evidence that letter identification recruits working memory under conditions of close spacing (Bacigalupo & Luck, 2015). Interestingly, we found that the spatial pattern of effects for working memory were opposite from the spatial pattern we observed for spelling, with visual working memory predicting lower crowding distances particularly for locations further from fixation and in the left visual field. This implies that there are distinct mechanisms for the effects of reading experience and visual working memory on crowding, and suggests that the intersection of visual working memory and crowding may be a fruitful area for future research. A recent letter crowding study with 50 observers (Kurzawski et al., 2023) also found high variability across individuals, and lower crowding thresholds in the right hemifield, similar to the present study. Such results suggest that to explain the high variability in crowding distances across observers and across the visual field, observer-level characteristics must be taken into account. Our results suggest that individual differences related to reading experience and visual working memory explain some of this variability.

The present evidence for experience-driven changes in the effects of crowding on letter identification extends previous evidence for a high degree of plasticity in the human visual system. Perceptual learning within the visual system has been observed in a wide array of laboratory tasks (Li, 2016), over longer term experience (Chopin, Bediou, & Bavelier, 2019), and in instances of disordered vision (Castaldi, Lunghi & Morrone, 2020). Other studies have shown effects of shorter-term experience on crowding for letters (Chung, 2007; Hussain et al., 2012) and on crowding for other learned, specialized visual stimulus classes such as musical notes (Wong & Wong, 2016). The present findings indicate that the effects of crowding are also influenced by a lifetime of reading experience.

CRediT authorship contribution statement

Kurt Winsler: Writing – review & editing, Writing – original draft, Visualization, Validation, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Steven J. Luck:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jml.2025.104689.

References

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods, 40*(1), 278–289.

Andrews, S., & Hersch, J. (2010). Lexical precision in skilled readers: Individual differences in masked neighbor priming. *Journal of Experimental Psychology: General, 139*(2), 299.

Andrews, S., Veldre, A., & Clarke, I. E. (2020). Measuring lexical quality: The role of spelling ability. *Behavior Research Methods, 52*, 2257–2282.

Bacigalupo, F., & Luck, S. J. (2015). The allocation of attention and working memory in visual crowding. *Journal of Cognitive Neuroscience, 27*(6), 1180–1193.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823.
Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*(5241), 177–178.
Brysbaert, M. (2019). How many words do we read per minute? a review and meta-analysis of reading rate. *Journal of Memory and Language, 109*, Article 104047.
Castaldi, E., Lunghi, C., & Morrone, M. C. (2020). Neuroplasticity in adult human visual cortex. *Neuroscience & Biobehavioral Reviews, 112*, 542–552.
Chakravarthi, R., Rubruck, J., Kipling, N., & Clarke, A. D. (2021). Characterizing the in-out asymmetry in visual crowding. *Journal of Vision, 21*(11), 10.
Chanceaux, M., & Grainger, J. (2012). Serial position effects in the identification of letters, digits, symbols, and shapes in peripheral vision. *Acta Psychologica, 141*(2), 149–158.
Chaney, W., Fischer, J., & Whitney, D. (2014). The hierarchical sparse selection model of visual crowding. *Frontiers in Integrative Neuroscience, 8*, 73.
Chopin, A., Bediou, B., & Bavelier, D. (2019). Altering perception: The case of action video gaming. *Current Opinion in Psychology, 29*, 168–173.
Chung, S. T. (2007). Learning to identify crowded letters: Does it improve reading speed? *Vision Research, 47*(25), 3150–3159.
Frömer, R., Dimigen, O., Niefind, F., Krause, N., Kliegl, R., & Sommer, W. (2015). Are individual differences in reading speed related to extrafoveal visual acuity and crowding? *PLoS One1, 10*(3), Article e0121986.
Grainger, J., Tydgat, I., & Isselé, J. (2010). Crowding affects letters and symbols differently. *Journal of Experimental Psychology: Human Perception and Performance, 36*(3), 673.
He, Y., Legge, G. E., & Yu, D. (2013). Sensory and cognitive influences on the training-related improvement of reading speed in peripheral vision. *Journal of Vision, 13 7: 14*, 1–14.
He, Y., & Legge, G. E. (2017). Linking crowding, visual span, and reading. *Journal of Vision, 17*(11), 11.
Hussain, Z., Webb, B. S., Astle, A. T., & McGraw, P. V. (2012). Perceptual learning reduces crowding in amblyopia and in the normal periphery. *Journal of Neuroscience, 32*(2), 474–480.
Johnson, M. K., McMahon, R. P., Robinson, B. M., Harvey, A. N., Hahn, B., Leonard, C. J., Luck, S. J., & Gold, J. M. (2013). The relationship between working memory capacity and broad measures of cognitive ability in healthy adults and people with schizophrenia. *Neuropsychology, 27*(2), 220–229. https://doi.org/10.1037/a0032060
Jones, J. P., & Palmer, L. A. (1987). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology, 58*(6), 1233–1258.
Joo, S. J., White, A. L., Strodtman, D. J., & Yeatman, J. D. (2018). Optimizing text for an individual's visual system: The contribution of visual crowding to reading difficulties. *Cortex, 103*, 291–301.
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26.
Kurzawski, J. W., Burchell, A., Thapa, D., Winawer, J., Majaj, N. J., & Pelli, D. G. (2021). The Bouma law accounts for crowding in fifty observers. *BioRxiv*, 2021.
Kwon, M., Legge, G. E., & Dubbels, B. R. (2007). Developmental changes in the visual span for reading. *Vision research, 47*(22), 2889–2900.
Kyllingsbaek, S., & Bundesen, C. (2009). Changing change detection: Improving the reliability of measures of visual short-term memory capacity. *Psychonomic Bulletin & Review, 16*, 1000–1010. https://doi.org/10.3758/PBR.16.6.1000
Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research, 48*(5), 635–654.
Levi, D. M., & Carney, T. (2009). Crowding in peripheral vision: Why bigger is better. *Current Biology, 19*(23), 1988–1993.
Li, W. (2016). Perceptual learning: Use-dependent cortical plasticity. *Annual Review of Vision Science, 2*, 109–130.
Louie, E. G., Bressler, D. W., & Whitney, D. (2007). Holistic crowding: Selective interference between configural representations of faces in crowded scenes. *Journal of Vision, 7*(2), 24.
Manassi, M., & Whitney, D. (2018). Multi-level crowding and the paradox of object recognition in clutter. *Current Biology, 28*(3), R127–R133.
McCarron, S. P., & Kuperman, V. (2021). Is the author recognition test a useful metric for native and non-native English speakers? An item response theory analysis. *Behavior Research Methods, 53*(5), 2226–2237.
Michael, G. A., & Ojéda, N. (2005). Visual field asymmetries in selective attention: Evidence from a modified search paradigm. *Neuroscience Letters, 388*(2), 65–70.
Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin, 137*, 267–296.
Moore, M., & Gordon, P. C. (2015). Reading ability and print exposure: Item response theory analysis of the author recognition test. *Behavior Research Methods, 47*, 1095–1109.
Nandy, A. S., & Tjan, B. S. (2012). Saccade-confounded image statistics explain visual crowding. *Nature Neuroscience, 15*(3), 463–469.
Paterson, K. B., McGowan, V. A., White, S. J., Malik, S., Abedipour, L., & Jordan, T. R. (2014). Reading direction and the central perceptual span in Urdu and English. *PLoS One1, 9*(2), Article e88358.
Petrov, Y., Popple, A. V., & McKee, S. P. (2007). Crowding and surround suppression: Not to be confused. *Journal of Vision, 7*(2), 12.
Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language, 14*(1), 174–180.
Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research, 46*(28), 4646–4674.
Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience, 11*(10), 1129–1135.
Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *Journal of Vision, 7*(2), 20.
Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language, 51*(2), 231–246.
Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506.
Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology. Human Perception and Performance, 37*, 514–528.
Ramamurthy, M., White, A. L., Chou, C., & Yeatman, J. D. (2021). Spatial attention in encoding letter combinations. *Scientific Reports, 11*(1), 1–12.
Risse, S. (2014). Effects of visual span on reading speed and parafoveal processing in eye movements during sentence reading. *Journal of Vision, 14*, 1–13.
Sagi, D. (2011). Perceptual learning in vision research. *Vision Research, 51*(13), 1552–1566. https://doi.org/10.1016/j.visres.2010.10.019
Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics, 74*, 5–35. https://doi.org/10.3758/s13414-011-0219-2
Snell, J. (2024). PONG: A computational model of visual word recognition through bihemispheric activation.*Psychological Review.* Advance online publication. doi: 10.1037/rev0000461.
Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, 402–433.
Tripathy, S. P., & Cavanagh, P. (2002). The extent of crowding in peripheral vision does not scale with target size. *Vision Research, 42*(20), 2357–2369.
Tydgat, I., & Grainger, J. (2009). Serial position effects in the identification of letters, digits, and symbols. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 480.
US Office of Management and Budget. (2016). September 30). Standards for maintaining, collecting, and presenting federal data on race and ethnicity. *Federal Register.* https://www.federalregister.gov/documents/2016/09/30/2016-23672/standards-for-maintaining-collecting-and-presenting-federal-data-on-race-and-ethnicity.
Vejnović, D., & Zdravković, S. (2015). Side flankers produce less crowding, but only for letters. *Cognition, 143*, 217–227.
Veldre, A., & Andrews, S. (2014). Lexical quality and eye movements: Individual differences in the perceptual span of skilled adult readers. *Quarterly Journal of Experimental Psychology, 67*(4), 703–727.
Veldre, A., & Andrews, S. (2016). Semantic preview benefit in English: Individual differences in the extraction and use of parafoveal semantic information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(6), 837.
Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*(2), 113–120.
Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences, 15*(4), 160–168.
Winsler, K., Grainger, J., & Holcomb, P. J. (2022). On letter-specific crowding and reading: Evidence from ERPs. *Neuropsychologia, 176*, Article 108396.
Wong, Y. K., & Wong, A. C. N. (2016). Music-reading training alleviates crowding with musical notation. *Journal of Vision, 16*(8), 15.
Yeatman, J. D., & White, A. L. (2021). Reading: The confluence of vision and language. *Annual Review of Vision Science, 7*, 487–517.
Zhao, C., Vogel, E., & Awh, E. (2022). Change localization: A highly reliable and sensitive measure of capacity in visual working memory. *Attention, Perception, & Psychophysics.*. https://doi.org/10.3758/s13414-022-02586-0