



# Crowding in depth for binocular and monocular observation

Lisa Valentina Eberhardt<sup>1</sup> · Anke Huckauf<sup>1</sup>

Published online: 18 March 2019  
© The Psychonomic Society, Inc. 2019

## Abstract

Crowding refers to the phenomenon of reduced recognition performance for peripherally presented targets that are flanked by similar stimuli. Crowding is known to vary with lateral distances (i.e., effects of target eccentricity and inter-character spacing). In the present experiment, we examined how crowding is affected by the distance of the stimuli in depth for natural viewing, i.e., for binocular observation of a real depth presentation. Superimposing the displays of two orthogonally arranged screens with a half-transparent mirror created real-depth presentation. We measured recognition performance of flanked compared to isolated targets that were presented at fixation depth, or in depths deviating from fixation depth (defocused). For both defocused directions (i.e., in front of and behind fixation depth), a near as well as a far distance from fixation was applied. Participants' task was to fixate a central cross at a constant distance (190 cm), and to indicate the gap position of an isolated or flanked Landolt ring that was presented at an eccentricity of 2°, on, in front of, or behind fixation depth. Results for natural binocular observation revealed increased crowding effects when stimuli were far compared to near from the fixation plane in depth. This resembles the common effect of eccentricity. Under monocular viewing, that is, without disparity information, crowding did not increase with increased depth distance. Thus, the result seemed to be an effect of binocular observation in real depth. This suggests that crowding in natural viewing might serve as a mechanism to stabilize and orient attention efficiently in three-dimensional space.

**Keywords** Crowding · Real depth · Depth perception · Natural viewing · Flanker interference

## Introduction

### Peripheral vision is mainly limited by crowding

Recognition of an isolated stimulus is limited in the periphery by a decline of visual acuity with increasing eccentricity. However, this decline is less severe than often suggested (Rosenholtz, 2016). In fact, peripheral vision is much more vulnerable to high stimulus density, an effect known as crowding. Crowding refers to the phenomenon of reduced recognition performance for a peripherally presented target stimulus in the presence of nearby flanker stimuli (Bouma, 1970). A demonstration of the effect with Landolt rings can be seen in Fig. 1. The crowding effect occurs for a broad range of stimuli (Levi, 2008; Pelli & Tillman, 2008; Whitney &

Levi, 2011), including letters (e.g., Huckauf & Heller, 2002), digits (e.g., Strasburger, 2005), symbols (e.g., Huckauf, Heller, & Nazir, 1999), faces (e.g., Farzin, Rivera, & Whitney, 2009), Gabor patches (e.g., Felisberti, Solomon, & Morgan, 2005), and even within complex objects (e.g., Martelli, Majaj, & Pelli, 2005, so-called internal crowding). The extent of the crowding effect across all different kinds of stimuli depends strongly on the spatial arrangement of target and flankers on the fronto-parallel plane (Pelli & Tillman, 2008): Crowding is modulated significantly by target-to-flanker spacing and by target eccentricity (e.g., Toet & Levi, 1992). The crowding effect increases when the target-to-flanker spacing becomes smaller, and also when the target's eccentricity increases.

Compared to visual acuity, the effect of eccentricity is considerably stronger for crowding (Bouma, 1970). Hence, crowding is an important limiting effect in peripheral vision (Rosenholtz, 2016; Strasburger, Rentschler, & Juettner, 2011). However, in natural vision, most of the objects seen at one glance do not appear at the fixation depth on a fronto-parallel plane. Rather, in natural vision objects are distributed across three-dimensional space, i.e. across real depth. Naturally,

✉ Lisa Valentina Eberhardt  
lisa.eberhardt@uni-ulm.de

<sup>1</sup> General Psychology, Institute of Psychology and Education, Faculty of Engineering, Computer Science and Psychology, Ulm University, Albert-Einstein-Allee 47, 89069 Ulm, Germany



**Fig. 1** Crowding effect. When fixating at the central fixation cross, crowding can be experienced for the flanked target: The leftward opened target Landolt ring is harder to recognize when presented with flankers (as in the right peripheral field) than in isolation (as in the left visual field)

these stimuli are observed binocularly. Thus, the question about crowding in natural viewing, which includes depth, arises. How is interference among adjacent stimuli pronounced when they are presented in depth?

### Crowding in depth

Although the spatial arrangement of stimuli was shown to be an important factor in crowding, the third spatial dimension, depth, has rarely been studied in crowding. Only a few studies have investigated the extent of crowding when the stimuli's depth was manipulated (Astle, McGovern, & McGraw, 2014; Eberhardt & Huckauf, 2017; Felisberti et al., 2005; Kooi, Levi, Tripathy, & Toet, 1994; Sayim, Westheimer, & Herzog, 2008). Most of them examine the assumption that differences between target and flanker depth (among other stimulus features like contrast or shape) reduce crowding (Astle et al., 2014; Felisberti et al., 2005; Kooi et al., 1994; Sayim et al., 2008). Therefore, depth was induced by binocular disparity to vary disparity between target and flankers. Thus, target and flankers were presented virtually in different depths (e.g., target in front of flankers). The results of those studies with stereoscopic depth indicate that crowding is reduced when targets and flankers differ in depth. This effect suggests that, analogously to the well-known effect of target-to-flanker spacing on the fronto-parallel plane (e.g., Toet & Levi, 1992), crowding is also reduced with increased target-to-flanker distance in depth.

In virtual depth using disparity as depth cue, however, stimuli, although appearing at various depths, are always presented on the same presentation plane (Hoffman, Girshick, Akeley, & Banks, 2008; Lamboojij, IJsselsteijn, Fortuin, & Heynderickx, 2009). Hence, observations from virtual depth are not simply transferable to real depth, that is, natural viewing conditions. Eberhardt and Huckauf (2017) describe how to examine crowding in real depth: With a real-depth presentation, target and adjacent flanker stimuli were always presented at the same depth, but deviated from the depth of fixation. Thus, subjects had to fixate at a certain distance while in front of or behind this fixation distance the flanked target stimulus was presented. The preliminary pilot data suggest that crowding effects in the investigated defocused depths ( $\pm .06$  dpt) do not differ from the fixated depth.

Understanding crowding in (real) depth requires diving a bit deeper into depth perception: Using a real-depth presentation avoids problems like lack of defocus blur, conflicting depth information, and vergence-accommodation mismatch, which are associated with stereoscopic depth presentation (Hoffman et al., 2008; Lamboojij et al., 2009). In real depth,

the eyes' vergence and accommodation are coupled, both focusing at the point of fixation. This physiological state of the eyes itself provides information about the absolute depth of fixation (Howard, 2012). Information about the relative depth distance of objects in front of or behind fixation depth is in real depth available by binocular disparity and by defocus blur (Howard, 2012). Binocular disparity refers to the difference in the images of an object in the two eyes, when this object deviates in depth from the fixation depth. Since binocular disparity is given as the vergence angle of the two eyes that would be required to fixate the defocused object, it depends on fixation distance and is more useful for smaller distances from the observer (Howard & Rogers, 2012).

Defocus blur refers to the increased blurriness of an object that is departing from the point of fixation either in front (i.e., it comes closer to the observer) or behind (i.e., receding from the observer; Howard, 2012). There is a certain range of tolerable defocus around the point of fixation, referred to as the depth of field (DOF), within which it is assumed that the level of blur is not detectable (Howard, 2012). Since the depth of field is given in diopters, its absolute size in object space depends on fixation distance, with blur being more informative for smaller distances from the observer. Taking into account empirical estimates showing a DOF of  $\pm 0.3$  dpt for a pupil diameter of 3 mm (e.g., Campbell, 1957) shows that defocused depths in Eberhardt and Huckauf (2017) were clearly within the DOF.

Thus considering both binocular disparity as well as defocus blur suggests that depth variations were probably too small to observe substantial differences between the tested depth conditions. Therefore, in the present study the range of distances was extended. Analogously to the effect of eccentricity, showing that crowding increases when stimuli are presented more distantly from the point of fixation (Levi, 2008; Pelli & Tillman, 2008; Whitney & Levi, 2011), for depth, one might assume that crowding increases when stimuli are presented farther apart from the fixated depth, e.g., all stimuli in front of or behind fixation depth.

### Present study

The aim of the present study was to investigate crowding for natural viewing. Therefore, the experimental approach of our former study (Eberhardt & Huckauf, 2017) was replicated and expanded by increasing the range of distances from fixation depth further in depth. Nevertheless, all defocused distances are still within the depth of field. To assess the role of natural viewing conditions further, a monocular control experiment

was conducted in order to distinguish effects of binocular disparity and potential effects of defocus blur.

Therefore, the same real three-dimensional experimental setup as in our former study (Eberhardt & Huckauf, 2017) was used to investigate crowding effects within the depth of field. In addition to defocused depths near to the fixated depth, two more depth planes farther away from fixated depth were applied. Thus, near and far depth distance (from fixation depth) was implemented, each in front of the fixated depth (i.e., between the observer and the fixation location) as well as behind the fixated depth (i.e., farther away from the observer than the fixation depth). All conditions were tested binocularly as well as monocularly to assess the role of natural binocular viewing.

For binocular observation, depth information should be available, especially by binocular disparity. Therefore, it was assumed that, similar to the effect of eccentricity, crowding effects increase among defocused depths with increased distance from fixation. Thus, crowding should be stronger in far compared to near distances. For monocular observation, if any, only defocus blur could be available as potential depth information. However, according to studies on defocus blur (e.g., Campbell, 1957) all defocused depths are within the DOF. Thus, defocus blur should be irrelevant for the current set of data. In this respect, one should assume crowding effects to be similar in far and near distances.

## Material and methods

### Participants

Sixteen (13 female;  $M_{age} = 21.5$  years,  $SD_{age} = 2.28$ ) psychology students of Ulm University participated in the Experiment. In a screening of binocular and monocular far visual acuity with a Landolt test chart, all participants passed the criterion of minimal visual acuity of 0.7 (decimal scale). To rule out persons with stereo disabilities, stereo acuity was tested using the TNO Test of Stereovision. All participants were able to identify at least objects at the level of the minimal criterion of 480 ArcSec, which is well above the smallest disparity presented during the experiment. The left eye was dominant in nine participants. The Experiment was conducted in four sessions on four different days. Prior to testing participants signed a written consent form that was in line with the guidelines of the German Research Foundation (DFG). They could receive partial course credit for participation.

### Apparatus

The experimental setup depicted in Fig. 2 was taken from Eberhardt and Huckauf (2017). Real depth was established by using a half-transparent mirror that was mounted in an

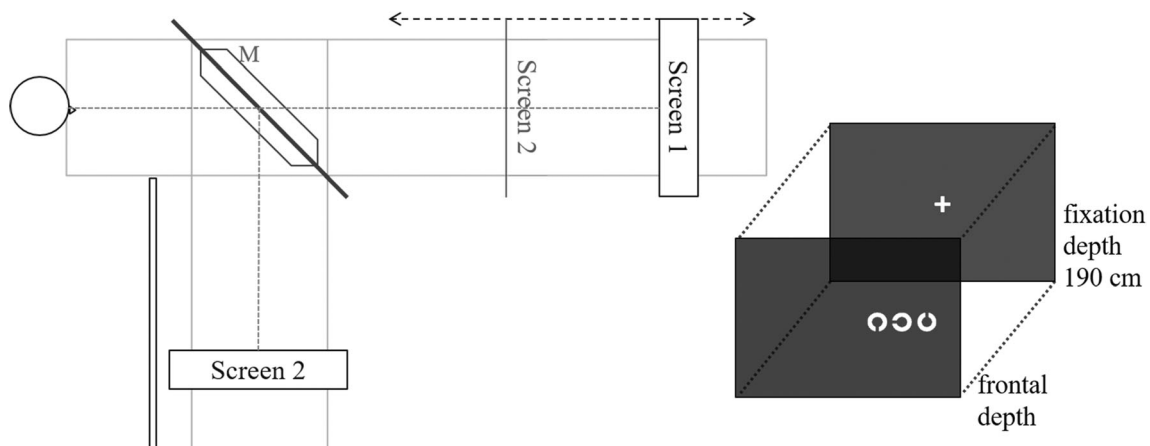
angle of  $45^\circ$  toward two orthogonally arranged screens. Stimuli were presented on the two simultaneously controlled 26-in. NEC MultiSync LCD screens (resolution  $1,440 \times 900$  px; refresh rate 60 Hz). The display of Screen 1 behind the half-transparent mirror was observed directly through the half-transparent mirror. Screen 2 was mounted orthogonal to the line of sight of the participant and in a  $45^\circ$  angle to the half-transparent mirror. Thus, the display of Screen 2 was reflected into the participant's line of sight. The distance between Screen 1 and the participant's eye position was adjustable to 150, 170, 215, and 240 cm. Screen 2 was fixed in a viewing distance of 190 cm. The fixation cross and all stimuli that appeared at fixation depth were presented on Screen 2. Thus, the display of Screen 1 was either in front of (150 and 170 cm) or behind (215 and 240 cm) fixation depth. Kooi, Dekker, van Ee, and Brouwer (2010) demonstrated that depth perception for briefly presented stimuli in such a real three-dimensional display is improved over a stereoscopic presentation with red-green anaglyphs.

The experimental program was controlled by MATLAB (Version 7.9) and the Psychtoolbox extension (Version 3; Brainard, 1997) on a Windows XP operating system with a Matrox M9138 LP graphics device.

### Stimuli and design

The fixation mark was a bright (white,  $160 \text{ cd/m}^2$ ) cross of  $0.6^\circ$  visual angle, centered on a dark background (black,  $0.2 \text{ cd/m}^2$ ). Stimuli were bright Landolt rings (white  $160 \text{ cd/m}^2$ ) with four possible opening directions (up-, right-, left-, and downward). Targets were displayed at  $2^\circ$  of eccentricity in the left and right visual field, either isolated or flanked to the left and right side. The center-to-center spacing of target and flankers was  $1^\circ$ . The flankers were randomly chosen under the constraint that the opening directions of the Landolt rings were incongruent with the target's and the other flanker's opening direction. Flankers were always presented in the same depth distance as the targets. The retinal stimulus size of targets and flankers was kept constant at  $0.6^\circ$  visual angle across all depth conditions. The near-depth distances (170, 215cm) constitute approximately a deviation of  $\pm 0.06$  dpt from fixation depth. The far distances (150, 240cm) are approximately  $\pm 0.1$  dpt. Data for near and far distances, each under monocular and binocular observation, were collected in four separate sessions. In each session 2 contexts (isolated, flanked)  $\times$  4 target ring openings (left, up, right, down)  $\times$  2 visual fields (left, right)  $\times$  3 depth directions (front, fixation, back) were repeated 20 times, resulting in 960 trials.

Taken together, the entire study consisted of four sessions on the basis of the binocular and the monocular observation condition and on the basis of the near- and the far-depth distances. The order of these four sessions was permuted on the basis of types of observation (binocular, monocular), distance



**Fig. 2** Left: Experimental setup. M indicates the half-transparent mirror that is mounted in an angle of  $45^\circ$  between the two orthogonally arranged screens. Screen 2 displays the fixation depth in a viewing distance of 190 cm for the reflected display, illustrated as the vertical line in the observer's line of sight. The distance of Screen 1 to the observer's eyes

was adjustable to 170, 215 ( $\pm 0.06$  dpt, near depth), 150, and 240 cm ( $\pm 0.1$  dpt, far depth). Right: Illustration of the displays of both screens when stimuli are presented defocused in front of the fixation depth (i.e., between observer and fixation depth)

(near, far), and the initial direction of the defocused depth plane (front, back) in Latin square. The resulting eight orders of sessions were balanced across the 16 participants.

## Procedure

To avoid confounding depth information by motion parallax, the participant's gaze was fixed by the use of a chin-rest. In the beginning of each session, the chin-rest was calibrated individually to assure that stimulus presentation was aligned. For monocular observation the participant's dominant eye was positioned at the central viewpoint, while the non-dominant eye was occluded. Prior to testing, in each session participants completed 72 training trials. The subsequent crowding experiment was split into two blocks. In one block the display of Screen 1 was presented in front of the fixation plane. In the other block, the display of Screen 1 was presented behind the fixation plane. Half of the trials for the fixation depth condition were presented in each block. Thus, each experimental block consisted of 480 trials.

The participants' task was to indicate the opening direction of the target Landolt ring. For response recording the number pad of a usual keyboard was used. The experimental procedure is illustrated in Fig. 3. Participants started a trial by pressing and holding a starting key (5 on the number pad). A trial started with the presentation of the fixation cross for 500 ms. Afterwards, stimuli were presented randomly in the left or right visual field for 20 ms to avoid saccadic eye movements toward the target (e.g., Robinson, 1964). A blank screen afterward ensured that stimuli remained in the iconic memory for processing (Sperling, 1960). In addition, within a presentation duration of 20 ms accommodative (e.g., Kasthurirangan & Glasser, 2006) or vergence movements (e.g., Bucci, Kapoula, Yang, & Bremond-Gignac, 2006) were impossible

(see also Dorman & van Ee, 2017). Participants were instructed to release the starting key and indicate the opening direction of the target Landolt ring as fast and as correctly as possible. The position of the response keys around the starting key corresponded to the four opening directions: Upper key to indicate an upward opening (number 8 on the number pad), right key for rightward opening, lower key for downward opening (2), and left key to indicate a leftward opening (4). In case there was no response within 1,000 ms after stimulus onset, an error sound rang out and the response was omitted. The next trial started by holding the starting key again.

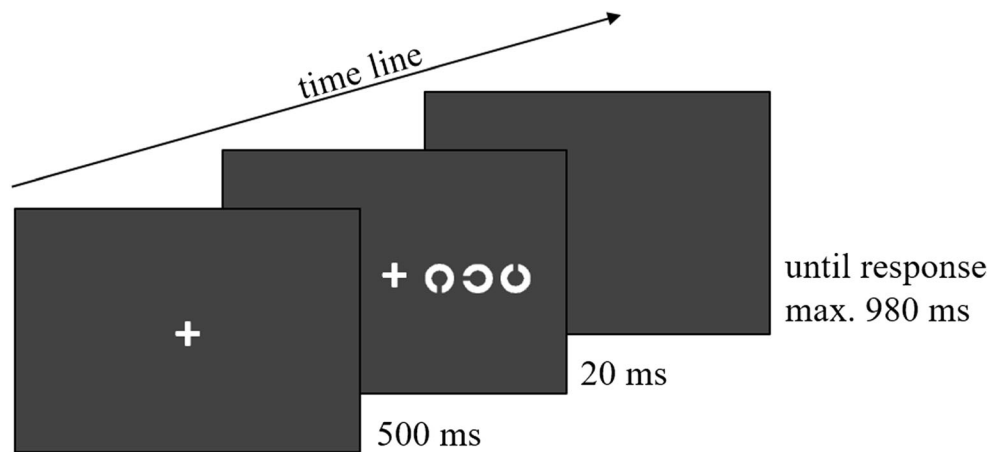
## Data analysis

Accuracy, defined as the proportion of correct target identifications, and reaction time for correct target identification were measured as dependent variables. Usually in crowding, accuracy-based measures are used (Bouma, 1970; Levi, 2008; Pelli & Tillman, 2008). However, both measures – accuracy and reaction time – determine recognition performance, but their sensitivity to specific perceptual processes can be different (e.g., Santee & Egeth, 1982). We therefore conducted our analyses using both measures. Statistical analyses were performed using IBM SPSS Version 24 (IBM Corp.). For statistical significance report an  $\alpha$ -level of  $p < .05$  was applied.

## Results

### Binocular observation

Table 1 shows descriptive data of the proportion of correct responses (accuracy, upper part of Table 1) and reaction time



**Fig. 3** Sequence of events within one trial. Each trial started with the presentation of the fixation cross for 500 ms. Then stimuli, in this example a flanked target, were presented for 20 ms. Participants had

maximally 1,000 ms after stimulus onset to respond according to the opening direction of the target Landolt ring, which is leftward in this example

(lower part of Table 1) for binocular observation. Inspection of descriptive data for isolated targets shows consistent performance across all depth conditions. This shows that all positions were within the DOF. Crowding effects were calculated for accuracy and reaction time by computing the difference in performance for flanked and isolated presentation. Crowding effects are plotted in Fig. 4.

**Accuracy**

For inferential analyses data for accuracy were transformed with arcsine using the formula  $F(x) = 2 \cdot \arcsine(\sqrt{x})$ . Further, crowding effects of arcsine-transformed data ( $M_{isolated} - M_{flanked}$ ) were referenced to fixation depth by subtracting crowding effects at fixation depth from crowding effects in each defocused depth. Thus, the difference in crowding effects referenced to fixation depth was used as the dependent variable for a  $2 \times 2$  repeated measures ANOVA with direction (front, back) and distance (near, far) of defocused depth as within-subject factors. The results revealed a significant main effect of distance,  $F(1,15) = 8.53, p$

$= .01, \eta_p^2 = .36$ , indicating larger crowding effects (referenced to fixation depth) in far,  $M = .15 (SE = .05)$ , compared to near distance,  $M = -.04 (SE = .03)$ . The main effect of direction and the interaction of direction and distance were non-significant,  $F(1,15) = .30, p = .59$  and  $F(1,15) = .22, p = .65$ , respectively.

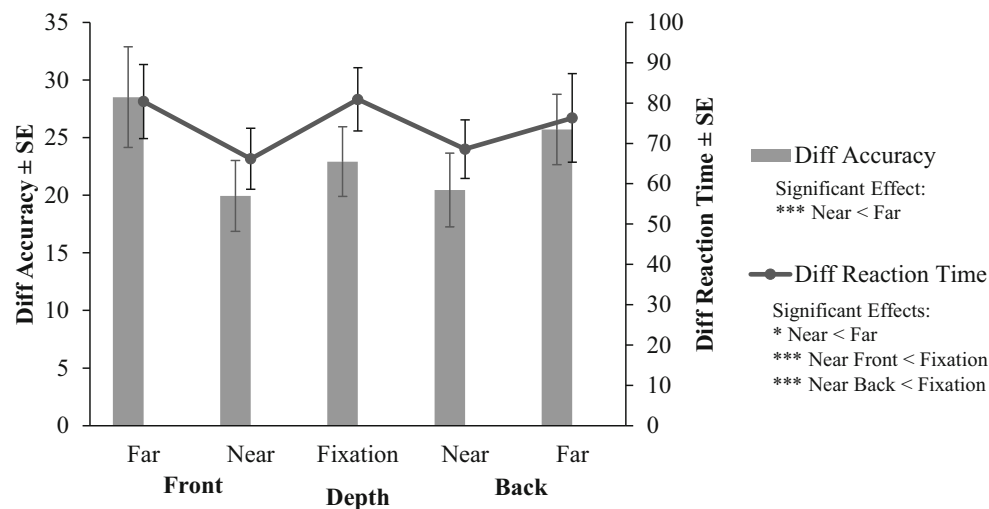
Whether crowding effects in defocused depths differed significantly from crowding effects at fixation depth was tested by Bonferroni-corrected one-sample t-tests. Therefore, the crowding effects, calculated as the difference between isolated and flanked conditions from arcsine-transformed data, were used as the dependent variable. None of the defocused conditions differed significantly from crowding effects at the fixated depth (all  $ps > .05$ ).

Thus, binocular accuracy shows that crowding effects differ systematically between defocused depths. The results indicate larger crowding effects in the far compared to the near depth distance. Since performance for isolated targets was homogenous, differences in crowding effects between depth conditions are mainly driven by the presence of flanking stimuli. This result resembles the effect of eccentricity on the fronto-parallel plane (e.g., Bouma, 1970; Huckauf et al., 1999), showing increased crowding with increased distance

**Table 1** Mean ( $M$ ) and standard error ( $SE$ ) of accuracy and reaction times as a function of context, direction, and distance for binocular observation

	Direction	Front		Fixation 0 $M (SE)$	Back	
		far $M (SE)$	near $M (SE)$		near $M (SE)$	far $M (SE)$
Accuracy (% correct)	isolated	95.48 (1.23)	96.24 (1.07)	95.55 (1.15)	96.93 (0.9)	96.69 (1.13)
	flanked	66.96 (4.81)	76.3 (3.72)	72.63 (3.6)	76.48 (3.66)	70.98 (3.48)
Reaction Time (ms)	isolated	582.87 (11.97)	559.92 (10.73)	558.27 (10.4)	560.78 (12.36)	570.74 (11.63)
	flanked	663.27 (15.03)	626.09 (10.03)	639.20 (11.3)	629.37 (11.04)	647.06 (15.73)

Note: Values for the fixation depth are averaged across both experimental sessions



**Fig. 4** Mean and standard error of crowding effect, defined as the difference in accuracy ( $M_{isolated} - M_{flanked}$ ), as well as reaction time ( $M_{flanked} - M_{isolated}$ ), is plotted as a function of depth for binocular observation. Note: \*  $p < .1$ , \*\*  $p < .05$ , \*\*\*  $p < .01$

between fixation and stimuli. Thus, interference among defocused stimuli became stronger when they were farther away from the fixated depth.

### Reaction times

For inferential analyses, crowding effects in reaction time ( $M_{flanked} - M_{isolated}$ ) were treated analogously to accuracy data. That is, they were also referenced to fixation depth by subtracting crowding effects at fixation depth from reaction time effects in each defocused depth condition. A  $2 \times 2$  repeated measures ANOVA with direction (front, back) and distance (near, far) of defocused depth as within-subject factors was conducted. Mirroring the results of accuracy, the analyses revealed a trend toward a significant main effect of distance,  $F(1,15) = 3.17$ ,  $p = .1$ ,  $\eta_p^2 = .17$ , pointing toward a higher deviance from fixation depth of the reaction time effect in near,  $M = -13.38$  ( $SE = 3.83$ ), compared to far distance,  $M = -2.23$  ( $SE = 3.92$ ). The main effect of direction and the interaction of direction and distance were non-significant,  $F(1,15) = .06$ ,  $p = .81$  and  $F(1, 15) = 1.6$ ,  $p = .22$ .

Whether crowding effects in defocused depths differed significantly from crowding effects at the fixated depth was tested by Bonferroni-corrected one-sample t-tests. Results indicated significantly less crowding than on the fixation depth for both near conditions (front and back),  $T(15) = 3.02$ ,  $p = .01$  and  $T(15) = 3.51$ ,  $p < .01$ , respectively.

Thus, corresponding to the results of binocular accuracy, crowding effects as measured by reaction time tended to be larger in far compared to near depths. Stronger interference among flanked stimuli in far distance from fixation is also reflected in reaction time effects.

### Monocular observation

Table 2 shows descriptive data of accuracy (upper part of Table 2) and reaction time (lower part of Table 2) for monocular observation. Inspection of descriptive data for isolated targets shows that stimulus presentation was also monocularly on a suprathreshold level. Crowding effects for monocular data were calculated for accuracy and reaction time as described for binocular data and plotted in Fig. 5.

### Accuracy

Again, for inferential analyses data for accuracy were transformed with arcsine and crowding effects were referenced to fixation depth. A  $2 \times 2$  repeated measures ANOVA with direction (front, back) and distance (near, far) of defocused depth as within-subject factors was conducted. The results showed no significant effect for direction, distance, or their interaction,  $F(1,15) = 2.18$ ,  $p = .16$ ,  $F(1,15) = .31$ ,  $p = .59$ , and  $F(1,15) = .12$ ,  $p = .73$ , respectively.

Whether crowding effects in defocused depths differed significantly from crowding effects at fixation depth was tested by Bonferroni-corrected one-sample t-tests. Again, the crowding effects calculated as the difference between isolated and flanked conditions from arcsine-transformed data were used as dependent variable. None of the defocused conditions differed significantly from crowding effects at fixation depth (all  $ps > .05$ ).

Hence, although crowding occurs in defocused depths, under monocular observation the extent of interference among stimuli does not depend on distance or direction of defocused depth.

**Table 2** Mean and standard error of accuracy and reaction time as a function of context, direction and distance for monocular observation

	Direction	Front		Fixation	Back	
		Distance	near		0	near
		far	near	0	near	far
		<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>	<i>M (SE)</i>
Accuracy (% correct)	isolated	94.85 (1.47)	94.84 (1.38)	93.26 (2.11)	93.94 (1.81)	93.62 (1.63)
	flanked	72.96 (4.01)	74.1 (4.01)	69.38 (3.89)	74.93 (4.05)	72.98 (4.2)
Reaction Time (ms)	isolated	584.59 (13.98)	573.18 (14.15)	577.81 (13.08)	582.12 (13.53)	598.96 (14.33)
	flanked	652.99 (16.53)	636.54 (14.65)	646.57 (14.11)	640.81 (14.28)	656.32 (14.7)

Note: Values for the fixation depth are averaged across both experimental sessions

**Reaction times**

For inferential analyses, again, reaction time effects referenced to fixation depth were used. A 2 × 2 repeated measures ANOVA with direction (front, back) and distance (near, far) of defocused depth as within-subject factors revealed only a significant effect of direction,  $F(1,15) = 7.79, p = .01, \eta_p^2 = .34$ , indicating larger differences from fixation depth for reaction time effects in the back,  $M = 10.74 (SE = 3.44)$ , compared to front,  $M = 2.89 (SE = 2.27)$ . The main effect of distance and the interaction of distance and direction were not significant,  $F(1,15) = 1.07, p = .32$ , and  $F(1,15) = 1.05, p = .32$ , respectively.

Whether crowding effects in defocused depths differed significantly from crowding effects at fixation depth was tested by Bonferroni-corrected one-sample t-tests. Results indicated only in the far back condition significantly less crowding than on fixation depth,  $T(15) = 3.83, p < .01$ .

Hence, reaction time effects did not differ between near- and far-depth distances, mirroring results of monocular recognition performance. However, beyond that, reaction time data indicates directional differences. Inspection of the descriptive

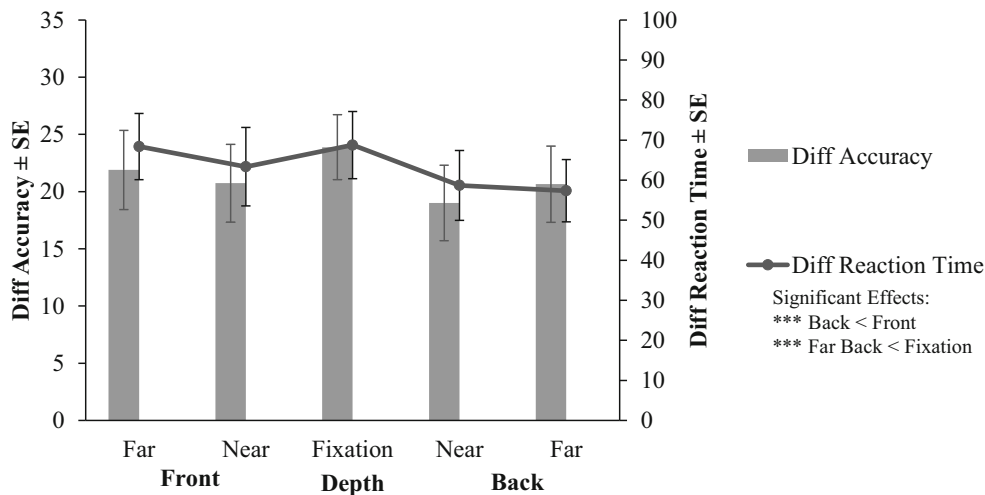
values in Table 2 suggests that this might be mainly due to increased reaction time toward isolated targets in the back.

**Discussion**

The main aim of the present study was to investigate crowding in natural viewing. Therefore, a real-depth presentation was used to examine crowding in defocused depths within the DOF. Isolated and flanked stimuli were defocused either in front of or behind fixation depth. In both of these directions a near and a far distance was tested. To assess the role of natural binocular viewing in crowding, a monocular control experiment was conducted. Thus, with monocular viewing, binocular disparity information is eliminated and potential effects of defocus blur are isolated.

**Crowding effects in real depth**

First, we consider natural viewing condition, that is, binocular observation. The comparison to the same depth condition (i.e., isolated and flanked targets both presented on fixation depth)



**Fig. 5** Mean and standard error of crowding effects for accuracy ( $M_{isolated} - M_{flanked}$ ) as well as reaction time ( $M_{flanked} - M_{isolated}$ ) are plotted as a function of depth for monocular observation. Note: \*  $p < .1$ , \*\*  $p < .05$ , \*\*\*  $p < .01$

reveals the following: In terms of reaction time there was a release from crowding in near defocused depths. However, in terms of accuracy, there was no difference to fixation depth. These findings replicate and extend the results of Eberhardt and Huckauf (2017), in which crowding for defocused stimuli was investigated in near-depth only. As in this previous study, in near-depth range crowding effects as measured by accuracy did not differ, but reaction time effects indicated less crowding for defocused stimuli.

Although crowding effects at fixation depth did not differ significantly from crowding in near distance, when inspecting the raw descriptive data of accuracy (instead of arcsine-transformed data) crowding effects at fixation depth even appear to be larger than in near distance. Also for monocular observation, the data show descriptively a similar increase of crowding effects at fixation depth compared to defocused depths. Hence, it seems that crowding effects differ slightly between the fixation depth and defocused depths. So, what characterizes trials that were presented at fixation depth? Stimuli presented on the same depth as fixation should be optically superior to defocused stimuli (e.g., in terms of contrast, blur) and thus are more salient (e.g., Artal, 2014). Since this concerns flankers to the same extent as targets, isolated target recognition as well as flanker interference might be increased at the fixation depth (Kothe & Regan, 1990; Simmers et al., 1999). However, we cannot exclude that this pattern is a confound of the experimental design. Since the number of trials on the fixation depth was split by experimental block, in each block two-thirds of trials were defocused stimuli while only one-third was on the same depth as fixation. Furthermore, same-depth trials were always presented on the reflected screen, while defocused stimuli were presented on the non-reflected screen. Taken together, future studies should clarify the relation between crowding at fixation depth and defocused depths.

For natural, that is binocular, viewing, the data indicate that crowding effects differ systematically between defocused depths: Crowding effects were stronger in the far- compared to the near-depth distances. Thus, analogously to the effect of eccentricity on the fronto-parallel plane (Bouma, 1970), crowding increases with increasing distance from fixation. This pattern was observed for binocular performance; in proportion of correct responses as well as by trend in reaction time data. However, as discussed, crowding tends to be larger at the fixation depth compared to near defocused depths.

Interestingly, the described difference in crowding effects between near and far depth from fixation did not occur under monocular observation. Neither accuracy nor reaction time data showed an effect of distance as found in the binocular data. This fosters the idea that the distance effect in the binocular data is due to the characteristics of disparity processing.

In general, comparing the data between accuracy and reaction time, the results mirror each other largely. Therefore, a

speed-accuracy trade-off (e.g., Fitts, 1966; Santee & Egeth, 1982) as a potential explanation for differences in crowding effects can be excluded. However, in the monocular data, reaction times revealed an additional effect, pointing toward a spatial asymmetry in depth: Correct reaction time toward isolated targets behind the fixation depth were heightened compared to reactions in front of the fixation depth. These results and the difference between binocular and monocular observation condition will be discussed in the next paragraphs on the basis of depth perception.

### Effects of defocus blur and disparity

The observed effects and differences between binocular and monocular observation become plausible when taking into consideration the constraints of depth perception with respect to the available sources of depth information. Since we used a real-depth presentation in the present study, we assume that vergence and accommodation were coupled and available as focus information. Further, in real depth binocular disparity and defocus blur are available as relative and ordinal depth cues, respectively.

The eye's vergence angle and state of accommodation should be coupled since we use a real-depth presentation (Lambooy et al., 2009). However, neither vergence (Mon-Williams & Tresilian, 1999) nor accommodation (Fisher & Ciuffreda, 1988) provide reliable egocentric distance information at the tested distances. However, it was shown that accurate fixation and focus of the eyes as given in real depth enhance perceived depth (Hoffman et al., 2008; Watt, Akeley, Ernst, & Banks, 2005). Thus, presentation of the to-be-identified target apart from the point of fixation (peripheral, either on the same depth plane, or on another depth plane) affected the perception of the target stimulus mostly insofar as relative binocular disparity and amount of defocus blur is concerned.

Nevertheless, it should be noted that the fixational state between binocular and monocular observation conditions differed. Stimulus presentation was kept physically identical, but for monocular observation, the eye position of the dominant eye was centered. Thus, vantage points between binocular and monocular observation differed. This means that exact eccentricity positions and disparity angles of stimuli differed between monocular and binocular observation. However, slightly differing vantage points between monocular and binocular observation should not affect relative distances and viewing angles between the different depth conditions within monocular and binocular observation. Thus, relative effects between depth conditions within both observation conditions should be still comparable. A more important difference between binocular and monocular observation is that binocular disparity information is missing in monocular vision.



Interestingly, in monocular viewing a directional effect was observed. A closer examination of the data revealed that this was mainly due higher reaction time for isolated targets behind compared to in front of fixation depth, indicating perceptual differences between the frontal and the retral direction. Also, Plewan and Rinkenauer (2017) have shown that simple reaction times toward closer targets are faster than toward targets that are farther away from the observer. Further, it is known that the monocular image is of lower contrast than the binocularly fused image (Blake & Wilson, 2011). Thus, optical aberrations that indicate the direction of depth (Howard, 2012) might have had a stronger impact in monocular than in binocular observation (Artal, 2014).

Defocus blur is usually regarded as an ordinal depth cue. Are increased crowding effects in far distance from fixation, as observed for binocular observation, due to increased blurriness of stimuli presented in far distances from fixation? A theoretical calculation of the depth of field indicates that our far distances are close to the borders of the depth of field (Green, Powers, & Banks, 1980). Furthermore, some studies suggest that blur perception is enhanced by the presence of nearby contours (Green et al., 1980), which would be the case in flanked conditions in our study. However, studies that measured the depth of field empirically with a variety of methods and stimuli all suggest that stimuli in our experiment are still within the depth of field (Marcos, Moreno, & Navarro, 1999; Yao, Lin, Huang, Chu, & Jiang, 2010). In addition, there is evidence that the depth of field increases in the periphery (Wang & Ciuffreda, 2004; Wang, Ciuffreda, & Irish, 2006). Moreover and most importantly, monocular results in the current study do not support the assumption that defocus blur drives increased crowding effects in far distance. Under monocular observation condition, if any, defocus blur could be available as depth cue (Vishwanath, 2012). However, we did not observe an increase in crowding in far distances as under binocular observation. Thus, the impact of defocus blur must be subordinate for the distances in the present study. In the tested range of depth, crowding seems to be unaffected by defocus blur.

For binocular observation, binocular disparity is a potent source of depth information for defocused stimuli in the present experimental setup. In particular, disparity is smaller for near compared to far distances from fixation. It is known that the retinal images on or close to fixation distance are fused. However, stimuli with larger disparities cannot be fused, and result in diplopia (Howard & Rogers, 2012). Even though it is known that the diplopia threshold increases towards the periphery, diplopia thresholds of stimuli at eccentricities up to 3° (outer flanker eccentricity in the present study) are almost as good as foveally (e.g., Blakemore, 1970; Mitchell, 1966; Vishwanath, 2012). Thus, it seems plausible that diplopia leads to reduced performance in far-depth distances, which was observed under binocular but not under monocular viewing condition. Failures during fusion of the three stimuli (i.e.,

the flanked target) result in overlaid images that impair recognition performance of the target. Interestingly, there was no drop in recognition performance for isolated targets in far distance. As Helmholtz (1867) already noticed, nearby stimuli reduce fusion limits (see also Howard & Rogers, 2012). Thus, fusion of flanked targets presumably resulted in more errors than for isolated targets. This could account for the drop in recognition performance for flanked but not for isolated targets, with increased depth distance from fixation. Hence, diplopia should be regarded as one mechanism underlying crowding effects in depth, which is also supported by monocular data. To quantify the extent that double images contribute to crowding in depth, in future studies diplopia should be measured in addition to crowding effects in real depth.

It is worth mentioning that the effects of depth were found for binocular observation, even though our stimuli were only briefly flashed. Studies have shown that it takes longer for apparent depth to emerge (e.g., Bradshaw, Hibbard, & Gillam, 2002; van Ee & Erkelens, 1996). However, binocular disparity is processed as soon as luminance contrast, that is, early on in visual processing (Caziot, Valsecchi, Gegenfurtner, & Backus, 2015). Thus, even though depth might not have been apparent yet, binocular disparity produced the described effects. Moreover, our sample was only screened for stereoability by using the TNO, which is based on red-green anaglyphs. Thus, the effect sizes in the current data might be diminished by inter-individual differences in stereoability (Dorman & van Ee, 2017; Kooi et al., 2010; van Ee & Richards, 2002) and could have benefited from a stricter criterion (Westheimer, 2013). For example, van Ee and Richards (2002) propose a more elaborate test of stereovision. Taken together, one might speculate that the effects could be even more pronounced when depth differences become apparent and when stereoability is extremely good in the sample. Future studies should address this issue.

### Crowding in natural viewing

The current results raise the question about the function of crowding in natural viewing. With respect to lateral space, crowding has often been regarded as a deleterious process (Levi, 2008; Whitney & Levi, 2011). However, it becomes more and more evident that crowding might be a helpful process in visual perception: For example, in the peripheral visual field, crowding might be regarded as a mechanism that results from an efficient processing of stimulus-dense regions in the periphery (Rosenholtz, 2017). One might speculate that efficient stimulus processing in the periphery, a possible cause of crowding, could even help to stabilize and orient information processing at the focus. Especially when taking into account the spatial specificity of crowding, this assumption might be regarded as plausible. The effect of stimulus eccentricity has been well known for a long time now: With increasing lateral

distance of stimuli from fixation, crowding increases. Furthermore, as the current study shows, crowding in defocused depth also increases with increasing distance of stimuli from the fixated depth (with the exception of the still unclear finding at fixation depth). Hence, one might assume that functionalities of crowding seem to apply to both the two-dimensional fronto-parallel plane as well as three-dimensional space: Increased crowding, that is stronger clutter, with increased distance from fixation (irrespective of whether laterally or in depth) might reduce distraction by a stimulus-rich environment and preserve capacity for information processing at the focus. However, whether similar mechanisms to two-dimensional space apply to three-dimensional space needs to be clarified.

## Conclusion

Taken together, our results indicate that under natural viewing conditions, that is, binocular observation in real depth, crowding effects increase with increased depth distance. However, the fixation depth seems to be an exception, since crowding here tended to be stronger compared to near defocused depth. Regarding situations when stimuli are presented defocused, one plausible speculation is that increased crowding in far depth serves as a mechanism to support and stabilize processes of selection in three-dimensional space. The monocular control experiment supports the idea that the effects in real depth are driven by conditions of natural binocular viewing, while defocus blur contributes less to the observed effects. Thus, increased crowding effects for natural viewing of stimuli in far defocused depth might be mainly due to binocular disparity, pointing toward double images as a potential mechanism of crowding in depth.

**Acknowledgements** We thank Gabriel Yuras for programming the experimental code and supporting data collection.

**Open Practices Statement** None of the data or materials for the experiments reported here are available, and none of the experiments were preregistered.

**Author contributions** The study concept and experimental design were developed by A. Huckauf and L. V. Eberhardt. Data collection, analysis, and interpretation were performed by L. V. Eberhardt under supervision of A. Huckauf. The manuscript was drafted by L. V. Eberhardt, taking into account critical revisions by A. Huckauf. Both authors approved the final version of the manuscript.

## References

- Artal, P. (2014). Optics of the eye and its impact in vision: A tutorial. *Advances in Optics and Photonics*, 6(3), 340. <https://doi.org/10.1364/AOP.6.000340>
- Astle, A. T., McGovern, D. P., & McGraw, P. V. (2014). Characterizing the role of disparity information in alleviating visual crowding. *Journal of Vision*, 14(6), 8. <https://doi.org/10.1167/14.6.8>
- Blake, R., & Wilson, H. (2011). Binocular vision. *Vision Research*, 51(7), 754–770. <https://doi.org/10.1016/j.visres.2010.10.009>
- Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. *The Journal of Physiology*, 211(3), 599–622. <https://doi.org/10.1113/jphysiol.1970.sp009296>
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, 226, 177–178.
- Bradshaw, M. F., Hibbard, P. B., & Gillam, B. (2002). Perceptual latencies to discriminate surface orientation in stereopsis. *Perception & Psychophysics*, 64(1), 32–40. <https://doi.org/10.3758/BF03194555>
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436. <https://doi.org/10.1163/156856897X00357>
- Bucci, M. P., Kapoula, Z., Yang, Q., & Bremond-Gignac, D. (2006). Latency of saccades, vergence, and combined movements in children with early onset convergent or divergent strabismus. *Vision Research*, 46(8–9), 1384–1392. <https://doi.org/10.1016/j.visres.2005.06.035>
- Campbell, F. W. (1957). The depth of field of the human eye. *Optica Acta: International Journal of Optics*, 4(4), 157–164. <https://doi.org/10.1080/713826091>
- Caziot, B., Valsecchi, M., Gegenfurtner, K. R., & Backus, B. T. (2015). Fast perception of binocular disparity. *Journal of Experimental Psychology: Human Perception and Performance*, 41(4), 909–916. <https://doi.org/10.1037/xhp0000066>
- Dorman, R., & van Ee, R. (2017). 50 years of stereoblindness: Reconciliation of a continuum of disparity detectors with blindness for disparity in near or far fepth. *I-Perception*, 8(6), 2041669517738542. <https://doi.org/10.1177/2041669517738542>
- Eberhardt, L. V., & Huckauf, A. (2017). Examining crowding using a real three-dimensional experimental setup. *Proceedings of the Latvian Academy of Sciences. Section B. Natural, Exact, and Applied Sciences.*, 71(5), 1. <https://doi.org/10.1515/prolas-2017-0055>
- Farzin, F., Rivera, S. M., & Whitney, D. (2009). Holistic crowding of mooney faces. *Journal of Vision*, 9(6), 18.1–15. <https://doi.org/10.1167/9.6.18>
- Felisberti, F. M., Solomon, J. A., & Morgan, M. J. (2005). The role of target salience in crowding. *Perception*, 34(7), 823–833. <https://doi.org/10.1068/p5206>
- Fisher, S. K., & Ciuffreda, K. J. (1988). Accommodation and apparent distance. *Perception*, 17(5), 609–621. <https://doi.org/10.1068/p170609>
- Fitts, P. M. (1966). Cognitive aspects of information processing: III. Set for speed versus accuracy. *Journal of Experimental Psychology*, 71(6), 849–857. <https://doi.org/10.1037/h0023232>
- Green, D. G., Powers, M. K., & Banks, M. S. (1980). Depth of focus, eye size and visual acuity. *Vision Research*, 20(10), 827–835. [https://doi.org/10.1016/0042-6989\(80\)90063-2](https://doi.org/10.1016/0042-6989(80)90063-2)
- Helmholtz, H. von. (1867). *Handbuch der physiologischen Optik*. Leipzig: Voss.
- Hoffman, D. M., Girshick, A. R., Akeley, K., & Banks, M. S. (2008). Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, 8(3), 33.1–30. <https://doi.org/10.1167/8.3.33>
- Howard, I. P. (2012). *Perceiving in depth: Volume 3: Other mechanisms of depth perception*. Oxford Psychology Series: Vol. 29. New York U. A.: Oxford Univ. Press.
- Howard, I. P., & Rogers, B. J. (2012). *Perceiving in depth: Volume 2: Stereoscopic vision*. Oxford Psychology Series: Vol. 29. New York U. A.: Oxford Univ. Press.
- Huckauf, A., & Heller, D. (2002). Spatial selection in peripheral letter recognition: In search of boundary conditions. *Acta Psychologica*, 111(1), 101–123. [https://doi.org/10.1016/S0001-6918\(02\)00045-8#](https://doi.org/10.1016/S0001-6918(02)00045-8#)

- Huckauf, A., Heller, D., & Nazir, T. A. (1999). Lateral masking: Limitations of the feature interaction account. *Perception & Psychophysics*, *61*(1), 177–189. <https://doi.org/10.3758/BF03211958>
- Kasthurirangan, S., & Glasser, A. (2006). Age related changes in the characteristics of the near pupil response. *Vision Research*, *46*(8–9), 1393–1403. <https://doi.org/10.1016/j.visres.2005.07.004>
- Kooi, F. L., Dekker, D., van Ee, R., & Brouwer, A.-M. (2010). Real 3D increases perceived depth over anaglyphs but does not cancel stereoaomaly. *Displays*, *31*(3), 132–138. <https://doi.org/10.1016/j.displa.2010.03.003>
- Kooi, F. L., Levi, D. M., Tripathy, S. P., & Toet, A. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision*, *8*(2), 255–279. <https://doi.org/10.1163/156856894X00350>
- Kothe, A. C., & Regan, D. (1990). Crowding depends on contrast. *Optometry and Vision Science : Official Publication of the American Academy of Optometry*, *67*(4), 283–286.
- Lambooi, M., IJsselstein, W., Fortuin, M., & Heynderickx, I. (2009). Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Technology*, *53*(3), 1–14. <https://doi.org/10.2352/J.ImagingSci.Technol.2009.53.3.030201>
- Levi, D. M. (2008). Crowding—an essential bottleneck for object recognition: a mini-review. *Vision Research*, *48*(5), 635–654. <https://doi.org/10.1016/j.visres.2007.12.009>
- Marcos, S., Moreno, E., & Navarro, R. (1999). The depth-of-field of the human eye from objective and subjective measurements. *Vision Research*, *39*(12), 2039–2049. [https://doi.org/10.1016/S0042-6989\(98\)00317-4](https://doi.org/10.1016/S0042-6989(98)00317-4)
- Martelli, M., Majaj, N. J., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision*, *5*(1), 58–70. <https://doi.org/10.1167/5.1.6>
- Mitchell, D. E. (1966). A review of the concept of "Panum's fusional areas". *American Journal of Optometry and Archives of American Academy of Optometry*, *43*(6), 387–401.
- Mon-Williams, M., & Tresilian, J. R. (1999). Some recent studies on the extraretinal contribution to distance perception. *Perception*, *28*(2), 167–181. <https://doi.org/10.1068/p2737>
- Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nat Neurosci*, *11*(10), 1129–1135. <https://doi.org/10.1038/nn.2187>
- Plewan, T., & Rinkenauer, G. (2017). Simple reaction time and size-distance integration in virtual 3D space. *Psychological Research*, *81*(3), 653–663. <https://doi.org/10.1007/s00426-016-0769-y>
- Robinson, D. A. (1964). The mechanics of human saccadic eye movement. *The Journal of Physiology*, *174*(2), 245–264. <https://doi.org/10.1113/jphysiol.1964.sp007485>
- Rosenholtz, R. (2016). Capabilities and limitations of peripheral vision. *Annual Review of Vision Science*, *2*, 437–457. <https://doi.org/10.1146/annurev-vision-082114-035733>
- Rosenholtz, R. (2017). Capacity limits and how the visual system copes with them. *Electronic Imaging*, *2017*(14), 8–23. <https://doi.org/10.2352/ISSN.2470-1173.2017.14.HVEI-111>
- Santee, J. L., & Egeth, H. E. (1982). Do reaction time and accuracy measure the same aspects of letter recognition? *Journal of Experimental Psychology: Human Perception and Performance*, *8*(4), 489–501. <https://doi.org/10.1037/0096-1523.8.4.489>
- Sayim, B., Westheimer, G., & Herzog, M. H. (2008). Contrast polarity, chromaticity, and stereoscopic depth modulate contextual interactions in vernier acuity. *Journal of Vision*, *8*(8), 12.1–9. <https://doi.org/10.1167/8.8.12>
- Simmers, A. J., Gray, L. S., McGraw, P. V., & Winn, B. (1999). Contour interaction for high and low contrast optotypes in normal and amblyopic observers. *Ophthalmic and Physiological Optics*, *19*(3), 253–260. <https://doi.org/10.1046/j.1475-1313.1999.00416.x>
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1–29. <https://doi.org/10.1037/h0093759>
- Strasburger, H. (2005). Unfocused spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision*, *5*(11), 1024–1037. <https://doi.org/10.1167/5.11.8>
- Strasburger, H., Rentschler, I., & Juetner, M. (2011). Peripheral vision and pattern recognition: a review. *Journal of Vision*, *11*(5), 13. <https://doi.org/10.1167/11.5.13>
- Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, *32*(7), 1349–1357.
- Van Ee, R., & Erkelens, C. J. (1996). Temporal aspects of binocular slant perception. *Vision Research*, *36*(1), 43–51. [https://doi.org/10.1016/0042-6989\(95\)00078-E](https://doi.org/10.1016/0042-6989(95)00078-E)
- Van Ee, R., & Richards, W. (2002). A planar and a volumetric test for stereoaomaly. *Perception*, *31*(1), 51–64. <https://doi.org/10.1068/p3303>
- Vishwanath, D. (2012). The utility of defocus blur in binocular depth perception. *I-Perception*, *3*(8), 541–546. <https://doi.org/10.1068/i0544ic>
- Wang, B., & Ciuffreda, K. J. (2004). Depth-of-focus of the human eye in the near retinal periphery. *Vision Research*, *44*(11), 1115–1125. <https://doi.org/10.1016/j.visres.2004.01.001>
- Wang, B., Ciuffreda, K. J., & Irish, T. (2006). Equiblur zones at the fovea and near retinal periphery. *Vision Research*, *46*(21), 3690–3698. <https://doi.org/10.1016/j.visres.2006.04.005>
- Watt, S. J., Akeley, K., Ernst, M. O., & Banks, M. S. (2005). Focus cues affect perceived depth. *Journal of Vision*, *5*(10), 834–862. <https://doi.org/10.1167/5.10.7>
- Westheimer, G. (2013). Clinical evaluation of stereopsis. *Vision Research*, *90*, 38–42. <https://doi.org/10.1016/j.visres.2012.10.005>
- 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. <https://doi.org/10.1016/j.tics.2011.02.005>
- Yao, P., Lin, H., Huang, J., Chu, R., & Jiang, B.-C. (2010). Objective depth-of-focus is different from subjective depth-of-focus and correlated with accommodative microfluctuations. *Vision Research*, *50*(13), 1266–1273. <https://doi.org/10.1016/j.visres.2010.04.011>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.