Perceptual integration across natural monocular regions

Katharina M. Zeiner
Modelling of Cognitive Processes, Technische Universität, Berlin, Germany
Vision Labs, School of Psychology and Neuroscience, University of St. Andrews, St. Andrews, Scotland

Manuel Spitschan
Vision Labs, School of Psychology and Neuroscience, University of St. Andrews, St. Andrews, Scotland
Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

Julie M. Harris
Vision Labs, School of Psychology and Neuroscience, University of St. Andrews, St. Andrews, Scotland

Natural scenes contain hidden regions, or occlusions, that differ in the two eyes, resulting in monocular regions that can only be seen by one eye. Such monocular regions appear to not be suppressed but seem to be integrated into the scene percept. Here we explore how the two eyes' views are combined to represent a scene that contains monocular regions, partially hidden behind a foreground occluding "fence." We measured performance in a density/numerosity discrimination task for scenes containing differing amounts of binocular and monocular information. We find that information from a number of separate monocular regions *can* be integrated into our overall percept of dot density/numerosity, although different observers use different strategies. If, however, both monocular and binocular information is present, observers appear to ignore the purely monocular regions, relying solely on the binocular information when making density/numerosity judgments. Our work suggests that binocular regions are favored over monocular regions, such that information from monocular regions is effectively ignored when binocular regions are present in a scene.

Introduction

The human visual system receives input from two forward-facing eyes that deliver significant overlap. A possible reason for this arrangement is to allow more of our cluttered world to be visible (Changizi & Shimojo, 2008). While this might have increased survival chances for our ancestors, this is still relevant today. When picking berries, there is a distinct benefit (less time spent near mosquitos) in detecting ripe berries quickly. Most of the time the berries are not in plain view but are partially "hidden" behind leaves and parts of the shrubs. As we will describe below, more of the berries are potentially visible if information from both eyes is used.

Under some circumstances, we know that the visual system is able to fill in information about a region that is not accessible, by interpolating information about the surface around that region (Durgin, Tripathy, & Levi, 1995). This is used to make an educated guess about what the inaccessible region is most likely to contain. Even though we have no visual input as to what these occluded regions contain, we perceive a coherent, complete object. This process is called amodal completion (Michotte, Thinès, Costall, & Butterworth, 1991; for a review see Sekuler & Murray, 2001).

Amodal completion of objects can slow performance during visual search tasks (e.g., He & Nakayama, 1992) and stereoacuity discrimination (e.g., Hou, Lu, Zhou, & Liu, 2006). It can improve texture segmentation (e.g., He & Nakayama, 1994) and pattern discrimination (e.g., Gold, Murray, Bennett, & Sekuler, 2000), and amodally completed objects can cause adaption (e.g., Fang & He, 2005). But, more importantly for the question of how information from monocular regions might be perceived, amodal completion highlights that there can be a marked dissociation between what is

Citation: Zeiner, K. M. Spitschan, M., & Harris, J. M. (2014). Perceptual integration across natural monocular regions. *Journal of Vision, 14*(3):5, 1–14. http://www.journalofvision.org/content/14/3/5, doi:10.1167/14.3.5.

doi: 10.1167/14.3.5 Received October 31, 2013; published March 5, 2014 ISSN 1534-7362 © 2014 ARVO

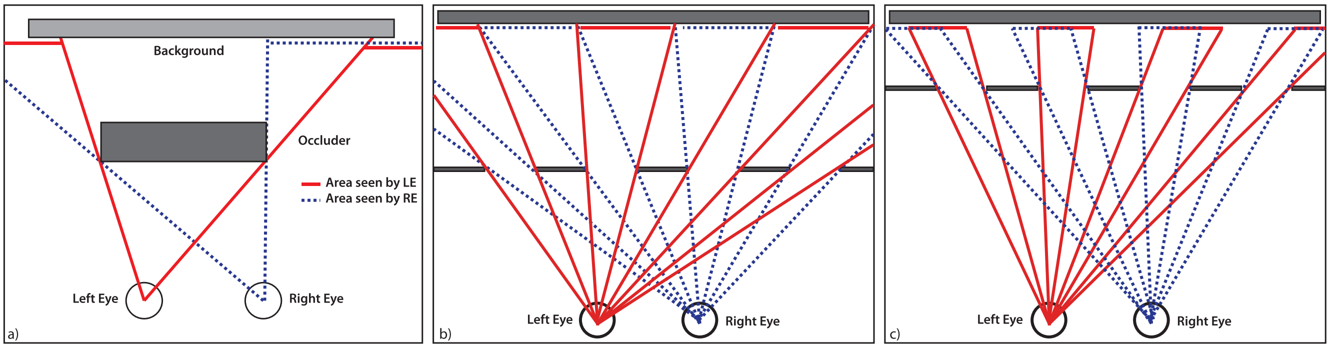

Figure 1. (a) Plan view of the eyes, an occluding object, and a background. Notice that there are regions of the background that only one eye can see. (b) Plan view of the arrangement first studied by Forte et al. (2002), showing the left and right eyes, a foreground occluding "fence" and background. Red lines show how the left eye sees only alternate strips of background and blue lines show how the right eye sees the other half of the background. The only binocularly visible regions in this display are the occluders in front of the background. (c) Plan view of the arrangement used in Experiment 2. Compared to the setup in (b), the occluders are moved closer to the background, thus leading to both monocular and binocular regions of the background being visible.

present in the retinal images and the percept we form based on this visual input.

Despite a wealth of literature on the effects of occlusion, almost all have considered it to be a 2D problem. In natural occlusion situations, however, occluded and occluding objects are often at different depths with respect to the observer. This results in differential occlusion between right and left eye (also referred to as da Vinci stereopsis, see e.g., Nakayama & Shimojo, 1990; or Harris & Wilcox, 2009 for a review). When one object occludes another, the foreground object lies closer than the background such that different parts of the background are occluded to the left and right eyes, respectively. This can be demonstrated by holding up a hand in front of the eyes and closing one eye. The viewing eye sees predominantly the hand, a region of the background is partly hidden, or occluded, by the hand. Through the other eye, one can see the hand, but also a different region of the background. This geometric configuration is illustrated in Figure 1a, which shows a pair of eyes, a foreground occluding object, and a background (viewed from above). The blue dotted line on the figure shows what regions of the background are visible to only the right eye: note the monocular zone to the right of the occluder that is only visible to this eye. The red line shows what parts of the background are only visible to the left eye. Most studies of amodal completion have used stimuli where the two eyes viewed identical scenes. As Figure 1a shows, this is not common under natural binocular viewing conditions. A very few studies have explored conditions like these. When the two eyes views are slightly different, consistent with a natural scene, amodal completion appears to be faster than when the two eyes views are identical (Bruno, Bertamini, & Domini, 1997). Further, it has been suggested that the perceived depth of background regions may be driven by amodal completion of monocular regions (Grove, Sachtler, & Gillam, 2006).

The existence of specifically monocularly occluded regions, at object boundaries, has been known about since Leonardo da Vinci (see Howard, 2012). More recently, it has been suggested that monocular regions themselves might provide a source of information about the depth between objects. This was coined "da Vinci stereopsis" by Nakayama & Shimojo (1990), who studied depth perception from monocular regions in detail (see also for example Gillam & Borsting, 1988; von Szily, 1921 [trans. by Ehrenstein & Gillam, 1998]; and for a recent review Harris & Wilcox, 2009). How the visual information contained within each monocular zone is represented has been much less explored.

Our aim here was to explicitly test how the visual system represents information in regions only visible to one eye. To do so we chose two experiments studying two types of occlusion geometries—one in which both monocular and binocular information was visible in a stimulus and the more extreme case in which only monocular information was visible.

Such an extreme occlusion situation occurs when a regular series of foreground objects (like a picket fence) causes all the visible parts of a background object to occur in monocular regions (Figure 1b). In this case the two eyes' views of the background scene are completely different. Notice how alternate regions of background are visible to just the left eye (red lines), or to just the right eye (blue lines). This viewing situation is akin to the more general phenomenon of binocular rivalry (e.g., Alais, O'Shea, Mesana-Alais, & Wilson, 2000; Blake, Lee, & Heeger, 2009; Blake & Logothetis, 2002). Rivalry occurs if the two eyes view totally different items, the resulting rivalrous percept switches between two different percepts. Yet for monocular regions like those in Figures 1a and 1b, rivalry typically does not occur (Forte, Peirce, & Lennie, 2002; Marlow, 2012). One study has demonstrated this by measuring contrast sensitivity when the two eyes view image patches containing different visual information. If the regions are surrounded by a binocularly visible contour, contrast thresholds within the patch are similar for each eye, suggesting that there is no suppression of one eye's view (Su, He, & Ooi, 2009). But, in that study, suppression did occur when the visible contour was removed. The question of how visual information is combined, or completed, across such regions has not been previously addressed.

We asked observers to decide whether a scene contained more elements than a comparison scene, and will refer to this as density/numerosity discrimination. There is a debate over whether this kind of task depends critically on the representation of density or of number (e.g., Burr & Ross, 2008; Durgin, 1995; Durgin, 2008; Sophian & Chu, 2008) or whether they are different at all (Dakin et al., 2011). We are agnostic over this issue; here, we simply chose the task for two reasons. First, because it requires integration of the available visual information across the whole scene. The second reason for choosing this task was that it has been suggested that density discrimination is represented at a level of processing where information has already been integrated between the two eyes (e.g., Durgin, 2001) while neuroimaging studies suggest that amodally completed objects may be represented very early (as early as V1) during visual processing (e.g., Rauschenberger, Liu, Slotnick, & Yantis, 2006).

In Experiment 1 we asked whether information from spatially distinct monocular regions was integrated across the two eyes, when there was no binocular information available. For this, we studied two extreme but natural occlusion situations, one containing monocular regions (caused by vertical occluders, as in Figure 1b), the other with binocular occlusions of parts of the background (caused by horizontal occluders).

We compared density/numerosity discrimination for three types of scene geometries: (a) a fully binocular scene where observers viewed a pattern of dots on a neutral gray background (Figure 2a); (b) a vertically occluded scene (as used by Forte et al., 2002; Figure 1b) in which each slice of background was only visible to one eye, but if the two eyes' views were combined all of the background was visible behind a foreground vertical slatted fence (Figure 2b); and (c) a horizontally occluded scene, where observers viewed a pattern of dots behind a foreground horizontally slatted fence (Figure 2c). Here, the fence and dotted background were both fully binocular, but half the dots were not visible to either eye as they were occluded by the fence in both eye's views.

In Experiment 2 we asked whether information in monocular regions is integrated with information in adjacent binocular regions to form a global density/numerosity percept. If the occluding fence in Figure 1b is moved closer to the background plane, then both monocular and binocular regions of the background become visible (Figure 1c). We studied density/numerosity discrimination for such scenes, varying the proportion of dots that was visible in monocular and binocular regions.

Experiment 1

Methods

Apparatus

Stimuli were presented on an Iiyama 22in Vision-Master-Pro monitor (resolution: 1280 × 1024 pixels, refresh rate: 100Hz; Visionmaster Pro, Tokyo, Japan) and viewed through a Modified Wheatstone Stereoscope. This was comprised of two sets of mirrors that are placed between a monitor and an observer so two images can be presented side by side on the screen, each visible to only one eye. The mirrors were aligned manually. The distance between the screen and the eyes was 100 cm. The head position was stabilized using a chin rest. Stimuli were generated and presented using MATLAB (MathWorks, Natick, MA) and the Psychophysics Toolbox 3 (Brainard, 1997; Pelli, 1997) on a PC workstation.

Participants

Eight participants, students aged 20–29, completed the study. All participants had normal or corrected to normal vision and normal binocular vision (TNO Stereo Test [Sussex Vision International, Rustington, UK] and Snellen EyeChart [Omega Healthcare, London, UK]). Participants were paid expenses for their participation. The experiment was approved by the University Teaching and Research Ethics Committee (UTREC) of the University of St Andrews. All participants gave written informed consent.

Stimuli

The baseline stimulus was a gray square of size 6.84° × 6.84° (luminance 32.47 cd/m^2), upon which a random pattern of black (luminance 0.01 cd/m^2) and white (luminance 66.54 cd/m^2) dots, of size 4.11 × 4.11 min arc, were superimposed. Half the dots were black, the other half white. Each dot had its polarity assigned randomly. Dot positions were also chosen randomly but no two dots could be located in the same position. The gray square was surrounded by a binary white noise texture (the luminance of each pixel, 1.37 × 1.37

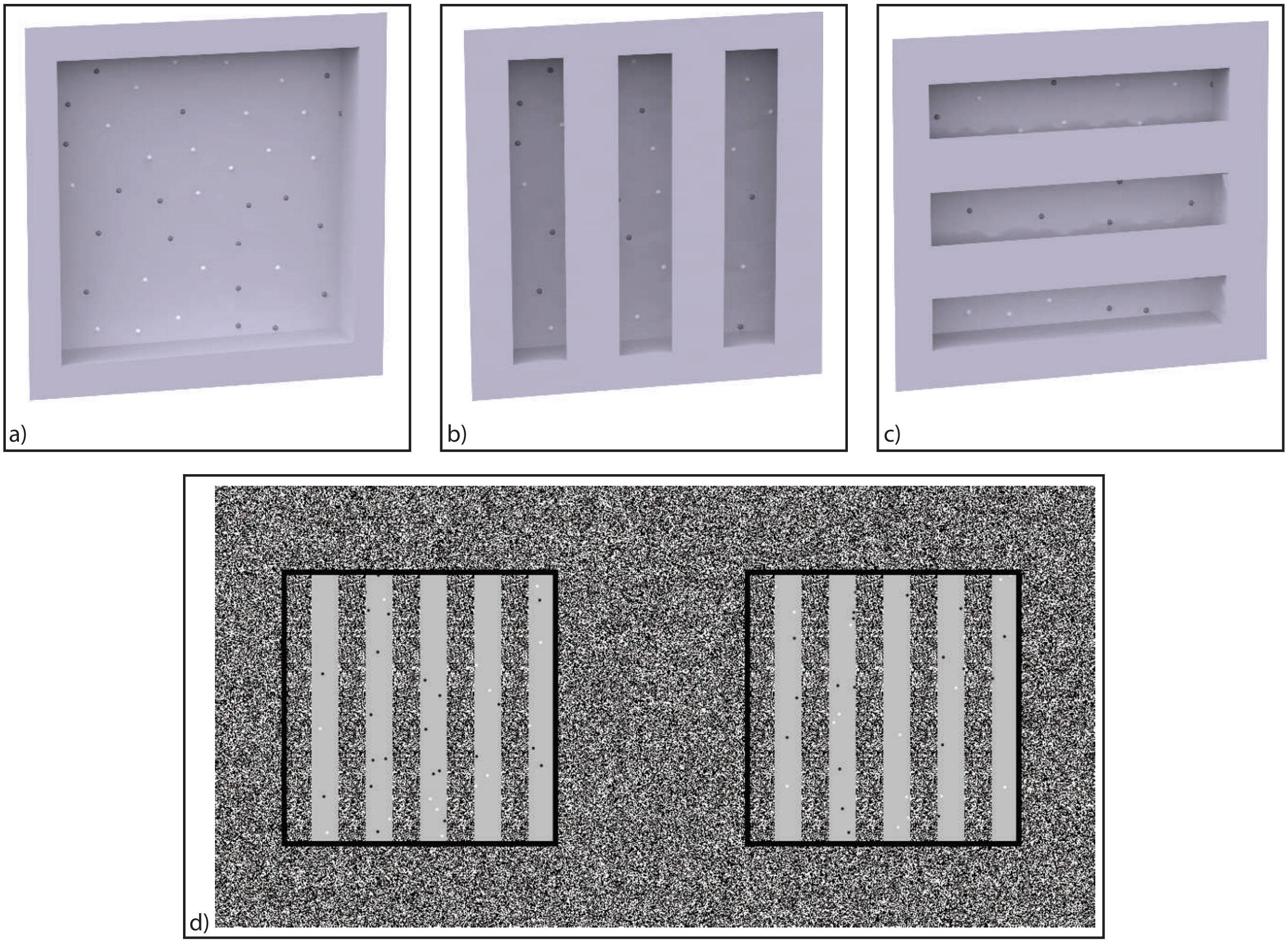

Figure 2. Cartoons of the stimuli used. (a) Fully binocular stimulus, (b) vertically occluded stimulus, (c) horizontally occluded stimulus, and (d) exampled stimulus showing left and right eye views (left eye view on left side, right eye view on right side).

min arc, was allocated randomly) that filled the remainder of the screen. The "standard" stimulus contained 80 dots (and was always an instance of the binocular stimulus 1 described below), while the "test" stimulus contained 45, 53, 64, 69, 80, 93, 100, 120, or 140 dots. We will express the difference between number of dots in the standard and test stimuli in terms of the proportion difference between them. This is defined as the difference between test and standard number, divided by the smaller of the two numbers (e.g., Van Oeffelen & Vos, 1982). Expressed in this way, we displayed a series of proportion differences of –0.75, –0,5, –0.25, –0.16, 0, 0.16, 0.25, 0.5, or 0.75.

We explored density/numerosity discrimination in three different types of test stimuli:

- Fully binocular: Both eyes viewed an identical pattern of dots (Figure 2a). The dots were located on a background gray plane displayed at a disparity of 41 min arc with respect to the plane of the screen, this corresponds to them being located 18.5 cm behind the white noise surround, presented in the plane of the screen.
- Vertically occluded: Observers viewed a foreground "fence" (binary white noise occluders, luminance: 34.80 cd/m^2, strip size: 0.68° $\times$ 6.84°), with strips of background (width: 0.68°, 41 min arc) only visible to one eye, or the other (Figure 2b). This was achieved by generating an identical background image for each eye, then shifting each eye's view of the background pattern of dots by 20.5 min arc away from the centre in opposite directions, to deliver a relative disparity of 41 min arc. All the dots in the background were visible, but all were visible to only one eye.
- Horizontally occluded: Both eyes viewed an identical pattern of random dots, half of which were hidden behind horizontally oriented foreground occluders (Figure 2c). The dots had a disparity of 41 min arc between the two eyes' views, the white noise surround and occluders were at zero disparity. The five horizontal occluders (binary white noise, as for

surround luminance: 34.69 cd/m^2) size: 6.84° × 0.68°) were spaced so that 41 min arc wide strips of the background pattern were visible between them.

To explore whether observers used primarily one eye, or the other, when viewing stimulus 2, we used different dot densities in specific regions of the display for all three stimuli. The background stimulus was divided into ten 41 min arc wide vertical strips. Half of the strips were assigned one-third of the dots; alternate strips were assigned the remaining two-thirds of the dots. For stimulus 2, this meant one eye was presented with only the lower-density strips, the other with only the higher-density strips. We presented the higher density to the right eye on 50% of trials. If observers used the eye with two-thirds of the dots and assumed the density in the background to be constant, we would expect the perceived density/numerosity of dots to be larger than for stimulus 1. For stimuli 1 and 3 the arrangement meant that the local density was different in different locations. This manipulation did not affect observers' ability to perform the task. They did not report a perceptual difference in the local densities. Figure 2d shows the stimulus as presented on-screen.

Procedure and task

In a two-interval forced choice (2IFC) task, participants were asked to indicate which of two intervals contained the stimulus with more dots.

Two stimuli were presented in each trial. The standard stimulus was always an instance of stimulus 1, the test stimulus could be any of the three stimuli. This delivered three conditions:

- Bin: Binocular—Binocular
- BinVer: Binocular—Vertically occluded (i.e., monocular regions)
- BinHor: Binocular—Horizontally occluded

Observers were presented with a fixation cross for 1 s, followed by the first stimulus interval (either test or standard), which was displayed for 0.4 s, a second fixation cross (1 s), and the second stimulus interval (0.4 s). Then a third fixation cross appeared, and stayed on the screen until participants made their response. Participants completed a total of 1,350 trials each (60 trials for each stimulus level for Bin and BinVer, 30 trials for each stimulus level for BinHor). Responses were given using one of two keys on a standard computer keyboard. In pilot experiments we found no difference in performance between 0.2 s presentation times and 0.4 s presentation times, but chose the longer presentation times because naive observers reported feeling quite rushed for the shorter presentation time.

Data analysis

We recorded the proportion of trials for which the test intervals were perceived as containing more dots, as a function of the proportion difference between test and standard. Psychometric functions were fitted with Psignifit 3.0 (Fründ, Haenel, & Wichmann, 2011) and had the general form:

$$\Psi(x;\theta) = \gamma + (1-\gamma-\lambda)F(x;\theta) \tag{1}$$

Where F is a sigmoid function, with a parameter vector $\theta = (\alpha,\beta,\gamma,\lambda)$. Alpha and beta describe the inflection point and the slope of the psychometric function, and gamma and lambda are the guess and lapse rate, respectively. Rather than constraining the possible values by setting a desired range, we "regularized" the parameters by applying so-called soft constraints. What this does is make certain values (the values we assume are most likely) more likely without imposing a stark cutoff point for the less likely values. Regularizing parameters has, in this context, the same effect as the assumed prior distribution in a Bayesian framework. For α a broad Gaussian distribution centered at 0 was assumed. For β a Gamma distribution was assumed. This means that the slope of the functions will be positive. In our case it makes sense to assume this because we expect participants to be more likely to respond that the test interval contained more dots when it actually contained more dots, and to respond that it contained fewer dots when the standard was more numerous.

Beta distributions were assumed for gamma and lambda to avoid some lapse/guess rates having a negative probability, a problem we would encounter with a Gaussian distribution for the priors and without having the steep cutoff we would encounter by using a uniform distribution. Constrained maximum likelihood estimation (MLE) was used to estimate the four parameters of the psychometric function. Confidence intervals were determined using a bias-corrected accelerated bootstrap (BCa; Efron, 1987).

The point of subjective equality (PSE) of the fitted function was defined as the proportion different that corresponds to 50% "test > standard" responses. Threshold was defined as the proportion difference between the 50% and 75% "test > standard" responses. Threshold and PSE estimates were obtained using constrained maximum likelihood estimation and the fitted functions resampled using the bootstrap method to retrieve confidence intervals for the estimates (Wichmann & Hill, 2001). To compare the threshold levels and PSEs between the different conditions, one-way ANOVAs were performed on the values obtained for thresholds and PSEs. The planned pairwise comparisons were Sidak corrected (e.g., Miller, 1981).

The data of all participants were used for further analyses. This means that, while there were individual

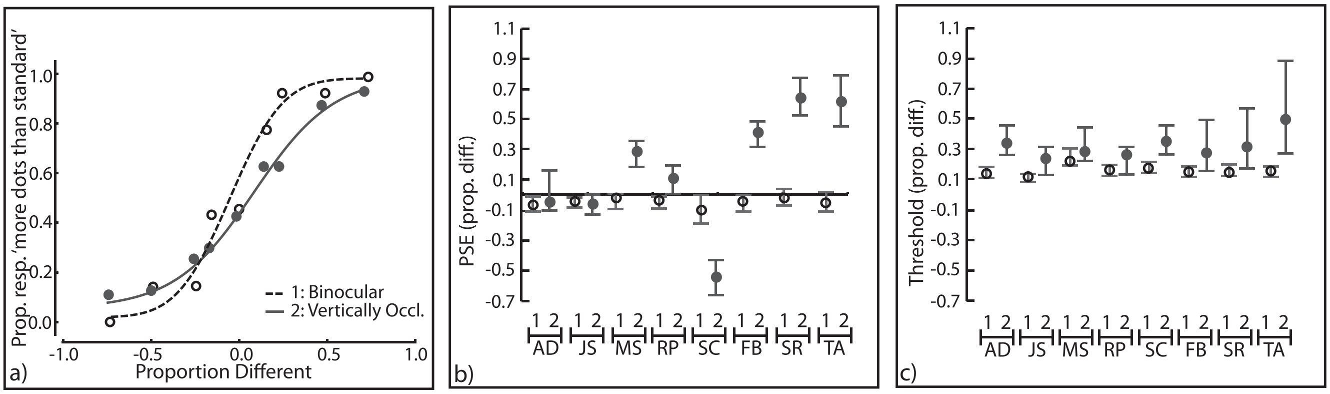

Figure 3. (a) Comparison of fitted sigmoids for stimuli 1 (black, dashed) and 2 (gray, solid) for participant RP. PSEs and thresholds of all participants are compared in (b) and (c), respectively. Error bars indicate 95% confidence intervals.

differences in participant performance, we were not preselecting for a specific behavior.

Results

One-way ANOVAs for the thresholds and PSEs indicated that participants showed significant differences in PSEs when comparing between the different conditions, $F(2, 21) = 4.13$, $p = 0.03$. There was also a significant difference in sensitivity between the different conditions, $F(2, 21) = 9.77$, $p < 0.01$.

Effects of monocular occlusion

Our main purpose here was to test whether the integration of information across several of these regions was as precise as when the observer had a binocular view.

Figure 3a shows example psychometric functions for Bin and BinVer, for observer RP. Figure 3c shows thresholds for all observers. The mean threshold for Bin was 0.17, and for BinVer it was 0.31. However, note the large overlap in the bootstrapped confidence intervals in Figure 3c for the two conditions, indicating no significant difference between thresholds in the two conditions. Within individuals, only three observers exhibited nonoverlapping confidence intervals. Thus, there is a clear (but nonsignificant) trend toward thresholds being higher for BinVer.

Figure 3b shows points of subjective equality (PSEs) from the psychometric functions. There was no overall significant difference between mean PSEs for the two conditions (Bin: PSE = –0.6; BinVer: PSE = 0.18; $p = 0.14$). However, Figure 3b demonstrates that there were large individual differences: five participants delivered strikingly different PSEs for the two conditions, with biases in different directions depending on the individual, and no overlap between confidence intervals. This suggests that different individuals integrate information in monocular regions in a variety of ways. What each might be doing will be discussed in detail below.

Is there evidence for suppression?

The vertically occluded condition was designed to deliver background dots to only one eye or the other, yet be consistent with a real-world situation in which the background was hidden behind a vertical slatted fence, in a different depth plane. The issue of interest is the extent to which the visual system can integrate information from across such a series of monocular regions. We tested for this by exploring any potential biases in the psychometric function. For example, the visual system might suppress one eye's view. As described in the methods section, one eye's view contained one-third of the dots the other, two-thirds. Had suppression due to rivalry occurred, the lower-density pattern would most likely have been suppressed (see Blake, Westendorf, & Overton, 1980).

As a control measure, we therefore analyzed BinVer separately for the situations when the higher density had been presented to the right eye and vice versa. We used the four parameters of the psychometric function for a repeated measures ANOVA. The interaction of the parameters with the eye of higher density were nonsignificant, $F(1.40, 19.66) = 0.08$, $p = 0.86$ (Greenhouse-Geisser corrected). There was also no significant main effect for the presentation eye, $F(1, 14) = 0.44$, $p = 0.84$. This suggests that, overall, observers do not use the high or low density eye differentially.

However, as described above, there were large individual differences. Almost all participants exhibited behavior that could be considered consistent with partially or fully ignoring one eye's view (irrespective of which eye is presented with the higher density) and filling in the missing information. This is visualized in Figure 4.

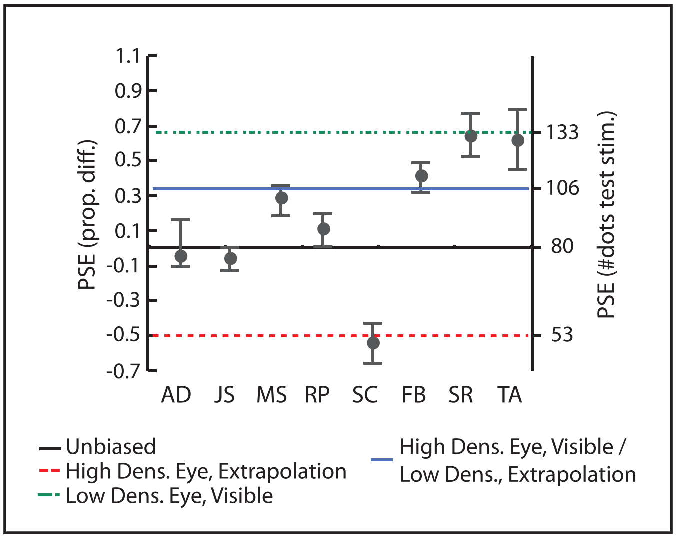

Figure 4. Comparison between PSEs for condition BinVer for all participants with predictions based on suppression of higher or lower density/numerosity eye's view. Error bars indicate 95% confidence intervals.

Unbiased performance (PSE of zero, when expressed as proportion different) would suggest that participants are using all the information present, across both eyes, and seamlessly integrating that information. Three observers exhibited this kind of behavior (AD, JS, and RP; black solid line in Figure 4). All other observers showed large biases.

There are several possible strategies to perform this task if the response is biased. If only one eye were used for viewing the vertically occluded stimulus, that eye views 50% dots and 50% occluder. Observers might respond purely to density, a behavior equivalent to assuming that as many dots were hidden behind the occluder as were visible, therefore assuming constant density in the background plane. Alternatively, observers might respond only to the number of dots visible, hence, with 50% hidden, they would respond as if the stimulus contained half as many dots as the binocular standard stimulus. Both of these strategies are considered below.

If participants responded only to the eye with higher density, and assumed the density of dots remained constant in the now occluded regions, they would overestimate the total number of dots by one-third. This means a vertically occluded stimulus would appear to contain approximately 107 dots compared with 80 in the binocular standard stimulus. For the two stimulus types to appear as if they contained the same number of dots, we would have to show 53 dots in the vertically occluded stimulus (i.e., one-third, or 27 fewer dots). This overestimation would shift the PSE to −0.5 (red, evenly dashed line in Figure 4). A bias of this magnitude was obtained for participant SC.

If participants used only the eye with higher density, and used only the visible dots, we would expect them to underestimate the density/numerosity of dots in the vertically occluded stimulus by one-third. If the two stimuli each contained 80 dots, a vertically occluded stimulus would appear to contain 53 dots. For the two stimulus types to appear as if they contained the same number of dots we would have to show 107 dots (i.e., one third, or 27 dots more) in the vertically occluded stimulus. If this underestimation occurred, it would shift the PSE to 0.33 (see blue, solid line, Figure 4). A bias of this magnitude was obtained for participants MS and FB (but see next paragraph below).

It is possible, though less likely, that the eye with lower density might be exclusively used. Using only the eye with lower density, and assuming constant density across the entire stimulus, we would expect observers to underestimate the density/numerosity of dots in BinVer by one-third. Note, that this is the same underestimation we would expect to see when using only the higher-density eye's view and only the visible dots. This underestimation would shift the PSE to 0.33 (see blue, solid line). Thus, using this analysis, we cannot distinguish between participants who rely on the higher-density eye and those who rely on the lower-density eye.

The final prediction of how observers could use one eye's view is also the most extreme: if they were to use only the lower-density eye's view and rely solely on the visible dots, we would expect observers to underestimate the density/numerosity of dots by two-thirds. This means 80 dots would be perceived as 27 dots, and to perceive 80 dots we would have to show 133 dots (two-thirds, or 53 dots more), leading to an expected bias of the PSE of 0.66 (green, unevenly dashed line). Participants SR and TA fall close to this prediction.

In sum, observers appear to use idiosyncratic strategies for "dot counting," suggesting many of them are not integrating information between the eyes, yet they deliver rather similar thresholds.

More general effects of occluders

So far we have compared a fully binocular stimulus with a vertically occluded one. This has not allowed us to measure the more general effect of introducing binocular occluders, which deliver spatially distinct binocular regions. It is possible that the 2D geometry of occluders could generate biases in the density/numerosity task. BinHor was designed to investigate this issue. Just like BinVer, the background plane was presented in spatially separated regions. Unlike in BinVer, these regions were all binocularly visible and only 50% of the background dots were visible through the occluders. Once again, there is the possibility that observers might respond to density (as if assuming half

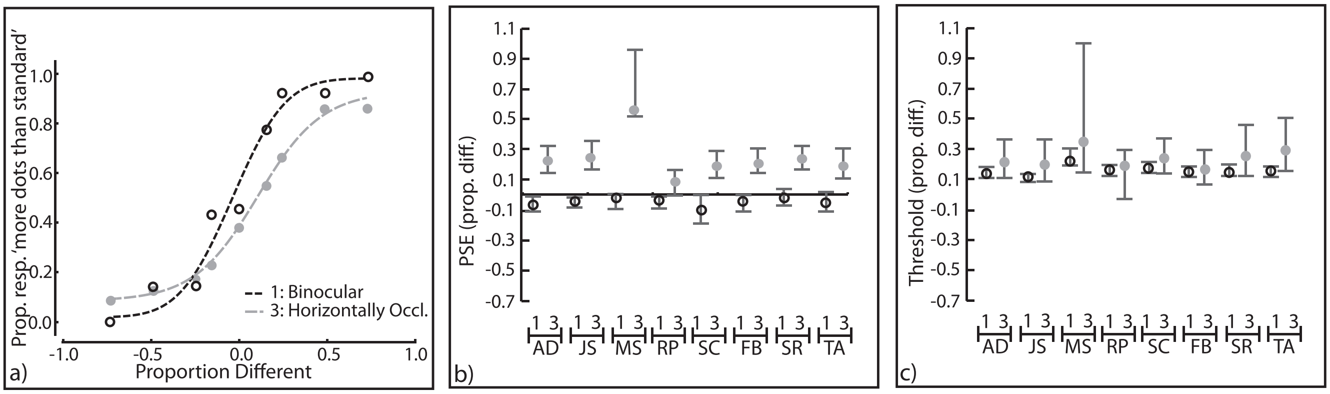

Figure 5. (a) Comparison of fitted sigmoids for stimuli 1 (black, evenly dashed) and 3 (gray, unevenly dashed) for participant RP. PSEs (50% point on the sigmoid) and thresholds (75% point on the sigmoid) of all participants are compared in (b) and (c), respectively. Error bars indicate 95% confidence intervals.

the dots were hidden), or they may respond to only those visible, in which case their responses will be biased, delivering an underestimation (positive PSE of 0.5).

Figure 5 compares the results for Bin and BinHor. Example psychometric functions are shown for one observer (Figure 5a); summary data showing PSEs (Figure 5b) and thresholds (Figure 5c) are shown for all observers.

PSEs were consistently lower for the binocular condition than the horizontal occlusion condition stimuli ($p = 0.02$, all differences between conditions reported here were calculated after the ANOVA described at the beginning of the results section; *p*-values were Sidak corrected to account for multiple comparisons). This demonstrates that the pattern presented in BinHor appears as less dense than the pattern presented in Bin. Note, however, that performance is not consistent with the prediction that observers count only visible dots (PSE = 0.5), except for one observer, MS, who also exhibited this behavior in BinVer, see Figure 4). While the average point estimates of the threshold for the two conditions were different (mean thresholds: Bin: 0.17 BinHor: 0.26; $p = 0.03$), the bootstrapped confidence intervals for the two conditions overlap for all participants (see Figure 5c). This indicates that the difference in sensitivity is not robust when individual observer behavior is considered.

Discussion: Experiment 1

We have found that there is an effect of monocular occlusion of background information on the way we process the density of random dot displays (comparing Bin and BinVer). However, this manipulation seems to predominantly affect accuracy and not sensitivity. When the stimulus is monocularly occluded, rather than binocular, large individual differences in PSE emerge, suggesting that individual's binocular integration of information is affected in different ways by the occlusion. Our vertically occluded stimuli delivered different sections of the background plane to the two eyes. We did not find that this affected observer sensitivity.

The biases in PSE suggest that binocular integration may not occur for five out of eight observers, with their behavior being consistent with using only information from one eye's view. However, three observers showed no bias, suggesting full binocular integration. These results therefore both compare and contrast with previous literature. For example, Forte and colleagues (2002) found that observers perceived a continuous coherent background surface composed of filtered noise texture presented in monocular regions. However, observers only achieved this percept when the background pattern was binocularly continuous and for specific spatial frequencies and orientations. Essentially, coherence was perceived when the pattern continued from one eye's monocular zone to the next. It is possible that observers did not integrate information in our study because our stimuli did not contain such continuous patterns.

Su and colleagues (2009) measured contrast sensitivity for a range of different stimuli, including some, like ours, where monocular regions were each surrounded by a binocular boundary. They found no reduction of sensitivity for such stimuli, suggesting that information could be integrated, rather than suppressed.

Previous literature does not explain the individual differences we found, particularly that some observers did seem to fully integrate across monocular regions, and others not at all. Recent work on numerosity and density perception (e.g., Dakin et al., 2011; Tibber, Greenwood, & Dakin, 2012) suggests that observers show a range of strategies when asked how many elements are present in a scene, some consistent with responding to density and some to number. Our data comparing a fully binocular view with a horizontally occluded but binocular view go some way toward exploring whether density or number are being used here. If observers are responding to numerosity then they will likely only respond to the visible dots in the horizontally occluded stimuli (data for MS is consistent with this hypothesis). If their percept is based on density then we expect they will be responding as if the number of dots remained constant behind the occluders (data for other observers were more consistent with this hypothesis). While one observer seems to be responding solely to the visible number of dots, all others appear to be responding more to density in the stimuli. No observer shows an unbiased percept of horizontally occluded stimuli. This suggests that that the density/numerosity extrapolation behind the occluders does not work seamlessly.

With this experiment we have shown that, while monocular regions in binocular scenes appear to be integrated into the overall percept of binocular scenes, observers seem to do so with different levels of success. The binocular regions in the scenes used here did not enable observers to segment the stimulus into foreground and background planes. Instead, most observers seemed to partially integrate information from the spatially distant monocular regions. Of course, the total monocular occlusion we used here would be very uncommon in the real world, where some binocular regions of background would be present in almost all situations. This was investigated further in Experiment 2, in which we asked whether observers are able to integrate monocularly and binocularly presented information not only to form a stable percept, but also to make judgments about the content of the regions.

Experiment 2

The aim of Experiment 2 was to investigate whether observers are able to integrate numerosities across adjacent monocular and binocular regions. More specifically, we asked two questions:

(a) Is the visual system able to integrate information across both monocular and binocular regions to obtain a stable representation of density or number? If this stimulus leads to a rivalrous percept if monocular and binocular regions are not integrated, this could show up in a heightened threshold for this condition (compared to the thresholds found in Experiment 1).

(b) Are monocularly and binocularly presented regions treated equally by the visual system, or does it rely more on one type of region than the other?

Methods

As in Experiment 1, observers viewed a background plane behind a set of occluders. The background plane and the occluder plane were spaced so that both monocular and binocular regions of the background were visible. The background plane consisted of a gray square of size 6.84 × 6.84 degrees (luminance 32.47 cd/m^2), which had black (luminance 0.01 cd/m^2) and white (luminance 66.54 cd/m^2) dots (4.11 × 4.11 min arc) distributed across it. Fifty percent of the dots were black, 50% white. The dots were arranged in "stripes" of dots that would be visible monocularly or binocularly. The remainder of the screen was filled by a binary white noise texture (the luminance of each pixel, 1.37 × 1.37 min arc, was allocated randomly). This was overlaid with fencelike vertical occluder stripes (made up of binary white noise, luminance: 34.80 cd/m^2, size: 0.68° × 6.84°). The occluders were spaced 0.68° (41 min arc) apart. While the occluders were placed in the same location in both eyes' views, the background plane was shifted outward by 0.23°, which meant that, between the occluders, there was a binocular region (width: 0.46°) as well as a monocular region (width: 0.23°) visible to each eye. This placement of the two stimulus planes meant the background plane appeared to lie at 6.18 cm behind the occluder plane. Figure 1c shows a schematic of the scene and how the monocular and binocular strips on the background were arranged. In this arrangement, there are regions of the background plane that are completely occluded from view from both eyes, regions that are occluded from view in one eye, and regions that are visible to both eyes.

As in Experiment 1, we developed stimuli in which a bias in density/numerosity perception would occur if observers were to use one stimulus region in preference to another. To achieve this, we altered the relative proportion of dots in the binocular and monocular regions.

Five different stimulus configurations were set up:

- Baseline: The monocular and binocular regions had the same density/numerosity of dots (i.e., one-half of dots were presented in binocular regions, one-half in monocular regions).
- Binocular High 1: Two-thirds of dots presented in binocular regions, one-third in monocular regions.
- Binocular High 2: Three-fourths of dots presented in binocular regions, one-fourth in monocular regions.
- Monocular High 1: One-third of dots presented in binocular regions, two-thirds in monocular regions.

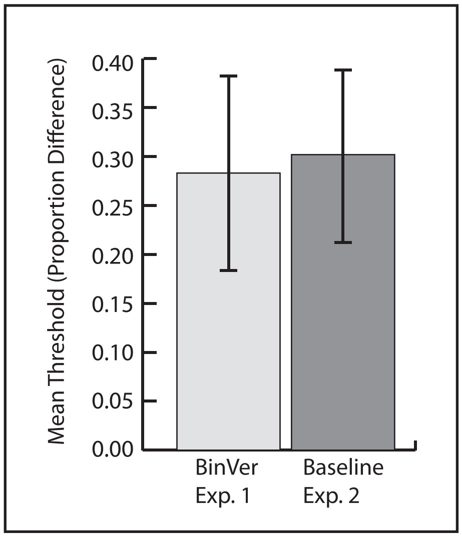

Figure 6. Comparison of vertically occluded conditions in Experiments 1 and 2. The error bars depict the 95% confidence intervals.

- Monocular High 2: One-fourth of dots presented in binocular regions, three-fourths in monocular regions.

Using a 2IFC task, participants were asked to indicate which one of two intervals contained the stimulus with more dots. Two stimuli were presented in each trial. The standard stimulus was always of type 1, the test stimulus could be any of the 5 stimulus configurations. This led to five experimental conditions:

- Baseline: Baseline—Baseline
- BinHigh1: Baseline—Binocular High 1
- BinHigh2: Baseline—Binocular High 2
- MonHigh1: Baseline—Monocular High 1
- MonHigh2: Baseline—Monocular High 2

If participants were to ignore the monocular regions and rely solely on the input from the binocular regions, we would arrive at two sets of extreme predictions. Consider the scenario in which participants base their density/numerosity judgment only on the binocularly visible dots. Here, the perceived density/numerosity of the Binocular High 1 stimulus would be overestimated by 16.7% compared to the presented baseline density/numerosity (66.7% of dots binocular compared to 50% binocular); the Binocular High 2 stimulus would have density/numerosity overestimated by 25%; the Monocular High 1 stimulus underestimated by 16.7%; and the Monocular High 2 stimulus underestimated by 25%.

In a second possible scenario, participants might ignore the monocular regions, sampling only from binocular regions, and assuming that the density of dots remains constant across the entire display. Under such circumstances, the perceived density/numerosity of the Binocular High 1 stimulus would be 33.3% higher than the presented density/numerosity; in the Binocular High 2 stimulus, the perceived density/numerosity would be overestimated by 50%; in the Monocular High 1 stimulus underestimated by 33.3%; and in the Monocular High 2 stimulus underestimated by 50%. Both sets of predictions are marked in Figures 7 and 8 so the biases can be compared to these predictions.

Four observers, students aged 22–25, participated in Experiment 2. Participants completed a total of 1,800 trials each (40 trials for each stimulus level for each condition). Responses were given using one of two keys on a standard computer keyboard. The data were analyzed in the same fashion as in Experiment 1.

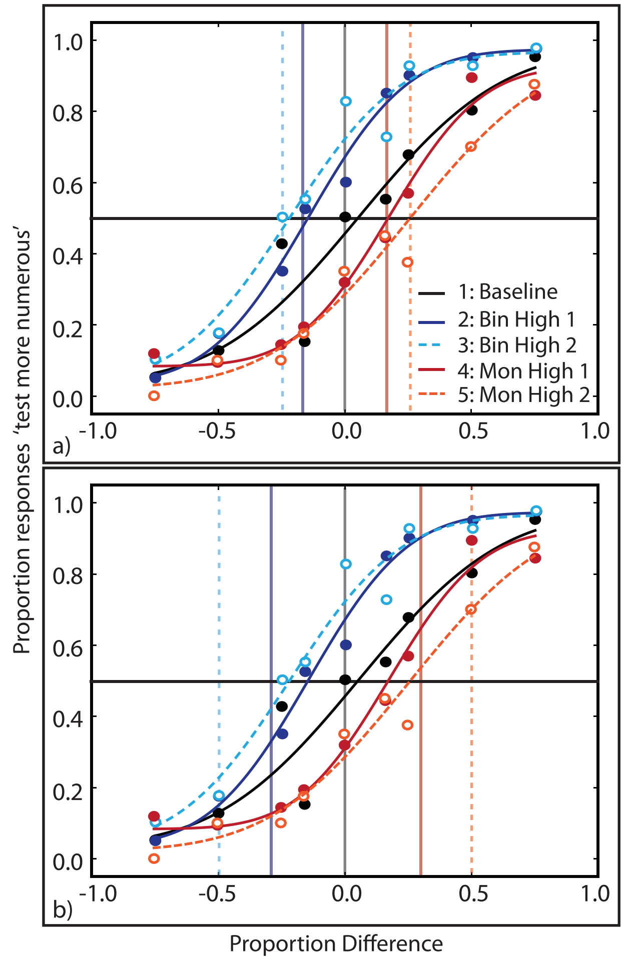

Figure 7. Fitted functions for participant BS for all conditions. The vertical lines in (a) are the predicted biases if the monocular regions were completely ignored and only the binocular regions compared across intervals. The vertical lines in (b) are the predicted biases if the monocular regions were completely ignored and a constant dot density was assumed across the entire background plane.

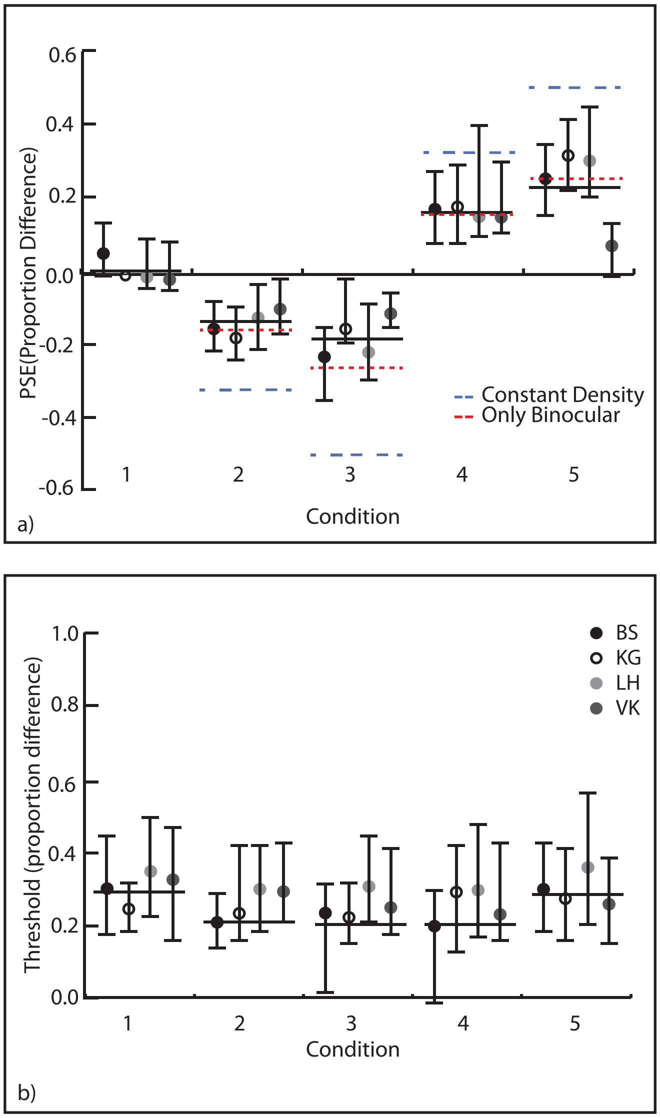

Figure 8. (a) The PSEs for the different conditions for all participants. Mean PSEs are displayed by the solid lines; predictions based on suppression of monocular regions are displayed by evenly (only binocular regions) and unevenly (constant density assumed) dashed lines. (b) The thresholds for the different conditions for all participants. Mean thresholds are displayed by the solid lines.

Results

As for Experiment 1, we were interested in observer thresholds and PSEs for the density/numerosity discrimination task. There was a significant difference between participants' PSEs across the different conditions, $F(4, 19) = 36.81$, $p < 0.01$. There was no significant difference in sensitivity between the different conditions, $F(4, 19) = 1.34$, $p = 0.30$. We will now consider the results in relation to the questions outlined above.

Are monocular and binocular regions integrated?

The main aim of Experiment 2 was to investigate whether monocular regions were as important as binocular regions for the representation of overall density/numerosity. We started by comparing sensitivity to the different stimuli across the two experiments. Participants reported a stable percept and were as sensitive to differences between two stimuli as participants were in Experiment 1. Figure 6 compares the vertically occluded conditions in the two experiments. BinVer from Experiment 1 delivered a mean threshold of 0.28. This is comparable to the threshold of 0.30 for the Baseline condition in the present experiment, where 50% of dots were in monocular regions, and 50% in binocular regions. When comparing between the five different conditions in Experiment 2, there was no significant difference ($p = 0.30$) between sensitivities (summarized in Figure 8b).

The effect of differing dot densities

Let us now consider how well participants perform when the density is different between the monocular and binocular regions. The aim of manipulating dot densities was to investigate whether the visual system is able to integrate information from both monocular and binocular regions. The baseline stimulus had a constant dot density across the stimulus plane. Participants could, in theory, completely ignore either monocular or binocular regions altogether in this baseline condition, and still show no bias in their performance. If a participant were to use this strategy for one of the "unbalanced" conditions, however, this would lead to a highly biased percept. Depending on how we expect binocular information to be used by the visual system we arrive at two possible predictions for observer responses. If monocular regions were completely ignored and the density/numerosity comparison was made based solely on the binocular regions (scenario 1). Alternatively, participants might also ignore monocular regions but assume a constant dot density across the entire display based on the density in binocular regions (scenario 2). Figure 7 shows the psychometric functions for participant BS. Figure 7a also shows the predicted PSE for scenario 1 (in which vertical lines cross the horizontal line at 0.5), and Figure 7b shows the predicted PSE under scenario 2 (in which vertical lines cross the horizontal line at 0.5).

First note that PSEs are different for the different conditions, revealing a bias. This demonstrates that monocular and binocular regions are not being treated equally. Biases were not as large as one would expect if a constant density of dots was assumed across the entire background plane (scenario 2; Figure 7b). The biases for this observer are more consistent with the scenario in which density/numerosity judgments are based solely on the binocular regions (scenario 1).

The remaining participants performed similarly, all showing considerably biased PSEs. Figure 8 shows the PSEs (Figure 8a) and thresholds (Figure 8b) for the different conditions for all participants. Black horizontal lines show the mean PSE across the four observers. The red horizontal lines in Figure 8a represent predictions for observer responses based on scenario 1 (evenly dashed, red line) or scenario 2 (unevenly dashed, blue line). Notice first how, for all conditions other than 1, PSEs are very different from the zero value. A PSE of zero is what we would expect if all the visual information presented is contributing to the density/numerosity task. The observed biases are a strong indicator that information presented in these monocular regions is not integrated into the overall percept when monocular regions are located immediately adjacent to binocularly visible regions of background plane. Observed PSEs for all participants were closer to the scenario 1 (red continuous dashed lines), suggesting that the visual systems of all participants used the binocular regions, and ignored the monocular regions of the background, when making the density/numerosity judgment.

Discussion: Experiment 2

The aim of Experiment 2 was to investigate whether and how well observers are able to integrate monocular and binocular information to form a stable percept and make judgments about its contents.

Our results demonstrate that, when we are required to integrate monocular information with binocular information, we rely solely on the binocularly presented information. This is potentially surprising, given that from Experiment 1 we know that the monocular information can be used when there are no binocular regions on the background plane. We also know that monocular information is used when making judgments about the two- and three-dimensional shape of an object, whether this object is visible (e.g., Wilcox & Lakra, 2007) or amodally completed (e.g., Bruno et al., 1997). Thresholds did not differ substantially across our five conditions (Figure 8), suggesting that the monocular regions do not interfere with processing of the binocular regions.

General discussion

In the two experiments presented here, observers had to judge the number of dots behind a set of occluders. While in the first experiment this required the integration of spatially distant monocular regions, in the second experiment this required integrating monocular regions with binocular regions adjacent to them. Observers appeared to integrate monocular regions in Experiment 1. However once they had to integrate across monocular and binocular regions (Experiment 2) a seemingly very different picture emerges. Here, they were unable to use monocular regions and instead suppress that information, relying solely on binocular information.

Our work compliments yet contrasts with that of the published literature. Forte and colleagues (2002) used a picket fence stimulus with the background containing only monocular regions, as we did in Experiment 1. Their participants were asked to rate the stability of the overall stimulus; in contrast, our experiments required participants to assess the content of the monocular regions and then integrate it to arrive at a density/numerosity judgment. Similar to Forte et al. (2002), we found that observers could do this, without cost to sensitivity, but with biases demonstrating individual differences in how the information was being used. Strikingly, when binocular information was additionally available in Experiment 2, monocular information seemed no longer useful at all. It appears as if participants suppress the monocular regions when binocular information about the same plane is present (here both monocular and binocular regions are located behind the occluders on the same background plane).

Our observers did not report a rivalrous percept in our experiments, and thresholds were not significantly different from those for fully binocular stimuli in Experiment 1 or Experiment 2. Arnold (2011) suggests there are several reasons why monocular regions do not cause binocular rivalry because there are differential occlusion cues for the two eyes (e.g., a monocular region will occur on the temporal side of the retinal image in relation to the occluder). For example, if you hold up a pencil between your eyes and either your computer screen or this page and read some text, you will be able to read the text while still being marginally aware of the pencil in front of it. Arnold (2011) argues this is achieved by active suppression of monocular regions. The findings in the two experiments presented here are in line with this suggestion. In Experiment 2, in which participants had to integrate monocular and binocular information to arrive at an unbiased percept, our data suggested integration of information from the two types of region did not occur. Participants appeared to consistently suppress the information provided in monocular regions.

In summary, we started with the suggestion (Changizi & Shimojo, 2008) that one reason for having binocular vision is to exploit the regions of background that only one eye can see, giving us the opportunity to sample more of the world than one eye would deliver.

Our first experiment suggested that this is possible, when the background region is completely monocular, observers can perform a density/numerosity task using information from those monocular regions. However, our second experiment demonstrated that, for the more general case where a background contains both binocular and monocular regions, the visual system behaves as if information from the monocular regions is not used at all. Our data are consistent with those regions being suppressed, and that only information viewed by both eyes is used for our density/numerosity judgment. We therefore provide evidence that must cast doubt on the generality of the appealing hypothesis that was first put forward by Changizi and Shimojo (2008).

Keywords: occlusion, monocular, binocular vision, representation

Acknowledgments

This study was supported by a Leverhulme Research Fellowship to Julie M. Harris; an Engineering and Physical Sciences Research Council (UK) (EPSRC) Studentship to Katharina M. Zeiner (2008–2011); EPSRC Grant EP/G038708/1 to Julie M. Harris; and KMZ was partly supported by German Research Foundation Grant DFG MA5127/1-1.

KMZ and JMH developed the study concept and contributed to the study design. Testing, experimental programming and data collection were performed by KMZ and MS. KMZ performed the data analysis and interpretation under the supervision of JMH. KMZ drafted the paper and JMH and MS provided critical revisions. All authors approved the final version of the paper for submission.

Commercial relationships: none.
Corresponding author: Julie M. Harris.
Email: jh81@st-andrews.ac.uk.
Address: Vision Labs, School of Psychology and Neuroscience, University of St. Andrews, St. Andrews, Scotland.

References

Alais, D., O'Shea, R. P., Mesana-Alais, C., & Wilson, I. G. (2000). On binocular alternation. *Perception*, *29*(12), 1437–1445.

Arnold, D. H. (2011). I agree: Binocular rivalry stimuli are common but rivalry is not. *Frontiers in Human Neuroscience*, *5*, 157, doi:10.3389/fnhum.2011.00157.

Blake, R., Lee, S., & Heeger, D. (2009). Binocular rivalry and neural dynamics. *Physical Review Letters*, *102*(25), 1–15.

Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, *3*, 13–21.

Blake, R., Westendorf, D. H., & Overton, R. (1980). What is suppressed during binocular rivalry? *Perception*, *9*(2), 223–231.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Bruno, N., Bertamini, M., & Domini, F. (1997). Amodal completion of partly occluded surfaces: Is there a mosaic stage? *Journal of Experimental Psychology: Human Perception and Performance*, *23*(5), 1412–1426.

Burr, D., & Ross, J. (2008). A visual sense of number. *Current Biology*, *18*(6), 425–428.

Changizi, M. A., & Shimojo, S. (2008). "'X-ray vision'" and the evolution of forward-facing eyes. *Journal of Theoretical Biology*, *254*, 756–767.

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences, USA*, *108*(49), 19552–19557.

Durgin, F. (1995). Texture density adaptation and the perceived numerosity and distribution of texture. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 149–169.

Durgin, F. H. (2001). Texture contrast aftereffects are monocular: Texture density aftereffects are binocular. *Vision Research*, *41*, 2619–2630.

Durgin, F. H. (2008). Texture density adaptation and visual number revisited. *Current Biology*, *18*(18), 855–856.

Durgin, F. H., Tripathy, S. P., & Levi, D. M. (1995). On the filling in of the visual blind spot: Some rules of thumb. *Perception*, *24*, 827–840.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, *82*, 171–200.

Ehrenstein, W. H., & Gillam, B. J. 1998. Early demonstrations of subjective contours, amodal completion, and depth from half-occlusions: Stereoscopic experiments with silhouettes by Adolf von Szily (1921). *Perception, 27*(12), 1407–1416.

Fang, F., & He, S. (2005). Viewer-centered object representation in the human visual system revealed by viewpoint aftereffects. *Neuron*, *45*(5), 793–800.

Forte, J., Peirce, J. W., & Lennie, P. (2002). Binocular integration of partially occluded surfaces. *Vision Research*, *42*, 1225–1235.

Fründ, I., Haenel, A. V., & Wichmann, F. A. (2011). Inference for psychometric functions in the presence of nonstationary behavior. *Journal of Vision*, *11*(6):16, 1–19, http://www.journalofvision.org/content/11/6/16, doi:10.1167/11.6.16. Retrieved from http://psignifit.sourceforge.net/. [PubMed] [Article]

Gillam, B., & Borsting, E. (1988). The role of monocular regions in stereoscopic displays. *Perception*, *17*, 603–608.

Gold, J. M., Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2000). Deriving behavioral receptive fields for visually completed contours. *Current Biology*, *10*(11), 663–666.

Grove, P. M., Sachtler, W. B., & Gillam, B. J. (2006). Amodal completion with background determines depth from monocular gap stereopsis. *Vision Research*, *46*, 3771–3774.

Harris, J. M., & Wilcox, L. M. (2009). The role of monocularly visible regions in depth and surface perception. *Vision Research, 49,* 1–20.

He, Z. J., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature, 359*, 231–233.

He, Z. J., & Nakayama, K. (1994). Perceiving textures: Beyond filtering. *Vision Research, 34*(2), 151–162.

Hou, F., Lu, H., Zhou, Y., & Liu, Z. (2006). Amodal completion impairs stereoacuity discrimination. *Vision Research*, *46*(13), 2061–2068.

Howard, I. P. (2012). *Perceiving in depth. Volume 3: Other mechanisms of depth perception*. Oxford, UK: Oxford University Press.

Marlow, P. (2012). Surface edges mitigate the disparity gradient constraint on binocular fusion and visual direction. *Vision Research*, *52*(1), 38–46.

Michotte, A., Thinès, G., Costall, A., & Butterworth, G. (1991). *Michotte's experimental phenomenology of perception*. Hillsdale, NJ: Erlbaum.

Miller, R. G. (1981). *Simultaneous statistical inference* (pp. 254–255; 2nd ed.). New York: Springer-Verlag.

Nakayama, K., & Shimojo, S. (1990). Da Vinci stereopsis: Depth and subjective occluding contours from unpaired image points. *Vision Research*, *30*(11), 1811–1825.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Rauschenberger, R., Liu, T., Slotnick, S. D., & Yantis, S. (2006). Temporally unfolding neural tepresentation of pictorial occlusion. *Psychological Science*, *17*(4), 358–364.

Sekuler, A. B., & Murray, R. F. (2001). Visual completion: A case study in grouping. In T. F. Shipley & P. J. Kellman (Eds.), *From fragments to objects: Segmentation and grouping in vision*, (pp. 265–294). New York: Elsevier.

Sophian, C., & Chu, Y. (2008). How do people apprehend large numerositites? *Cognition*, *107*(2), 460–478.

Su, Y., He, Z., & Ooi, T. L. (2009). Coexistence of binocular integration and suppression determined by surface border information. *Proceedings of the National Academy of Science, USA*, *106*(37), 15990–15995.

Tibber, M. S., Greenwood, J. A., & Dakin, S. C. (2012). Number and density discrimination rely on a common metric: Similar psychophysical effects of size, contrast, and divided attention. *Journal of Vision*, *12*(6):8, 1–19, http://www.journalofvision.org/content/12/6/8, doi:10.1167/12.6.8. [PubMed] [Article]

Van Oeffelen, M. P., & Vos, P. G. (1982). A probabilistic model for the discrimination of visual number. *Perception & Psychophysics*, *32*(2), 163–170.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, *63*(8), 1293–1313.

Wilcox, L. M., & Lakra, D. C. (2007). Depth from binocular half-occlusions in stereoscopic images of natural scenes. *Perception*, *36*(6), 830–839.