Binocular disparity augments the capacity of vision without affecting subjective experience of depth

Henry Railo* [1,2], Joni Saastamoinen[2], Sipi Kylmälä[2], and Aapo Peltola[3]

[1] Department of Clinical Neurophysiology, University of Turku and Turku University Hospital, FI-20521, Turku, Finland

[2] Department of Psychology, University of Turku, 20014, Finland

[3] Turku Game Lab, Turku University of Applied Sciences, 20520, Finland

* Corresponding author:
Henry Railo,
Department of Psychology,
University of Turku,
Finland
Email: hmrail@utu.fi
Phone: +358-45-6765227

Binocular disparity results in a tangible subjective experience of three-dimensional world, but whether disparity also augments objective perceptual performance remains debated. We hypothesized that the improved coding of depth enabled by binocular disparity allows participants to individuate more object at a glance as the objects can be more efficiently differentiated from each other and the background. We asked participants to enumerate objects in briefly presented naturalistic (Experiment 1) and artificial (Experiment 2) scenes in immersive virtual reality. This type of enumeration task yields well-documented capacity limits where up to 3-4 items can be enumerated rapidly and accurately, known as subitizing. Our results show that although binocular disparity did not yield a large general improvement in enumeration accuracy or reaction times, it improved participants' ability to process the items right after the limit of perceptual capacity. Binocular disparity also sped-up response times by 27 ms on average when artificial stimuli (cubes) were used. Interestingly, when artificial stimuli were used, the perceptual advantage provided by binocular disparity emerged without a concurrent enhancement in subjectively experienced depth. This suggests that the functional and subjective sides of stereopsis can be dissociated. Altogether our results suggest that binocular disparity may increase the number of items the visual system can simultaneously process. This may help animals to better resolve and track objects in complex, cluttered visual environments.

Keywords: binocular disparity, stereopsis, 3D vision, virtual reality, subitizing

Humans and other primates have two forward facing eyes and a large overlap between their visual fields (Heesy, 2009). One of the major benefits of this is stereopsis: integration of the slightly different viewpoints of the two eyes (binocular disparity) results in tangible subjective sensation of three-dimensional visual space, and accurate coding of depth (Barry, 2009; Vishwanath, 2014). Binocular disparity improves eye-hand coordination (Melmoth, et al., 2009; O'Connor, et al., 2010; Fielder & Moseley, 1996), and visual recognition of singular items when simple stimulus displays are used (Edelman & Blüthoff, 1992; Caziot & Backus, 2015; Lee & Saunders, 2011; Wardle, et al., 2010). Whether disparity also improves the segmentation of objects from background in naturalistic settings remains open (Valsecchi, et al., 2013). Specifically, we were interested in examining if binocular disparity improves the recognition of multiple objects in naturalistic settings. We hypothesized that because disparity allows differentiating between different objects in depth it enables faster and more accurate individuation of multiple objects. This type of perceptual enhancement could enable more efficient visually-guided behavior, and hence improve the individual's chances of survival.

Stereopsis is one of the defining features of primates (Ross & Martin, 2006), and linked to the expansion of brain structures that process visual information (Barton, 2004). What benefits drove the emergence of frontal eyes during primate evolution? Early theories explained the emergence of frontal eyes as adaptations to arboreal locomotion because movement in arboreal environment (for instance, when leaping between tree branches) requires accurate judgement of distances (Jones, 1917; Collins, 1921; Smith, 1924). The arboreal theory of stereopsis has been challenged as, for instance, many non-primate species (such as the squirrel) are capable of skilled arboreal locomotion despite the lack of a binocular visual field (Cartmill, 1974) suggesting that arboreal locomotion alone does not explain the evolution of frontal eyes. According to the alternative visual predation hypothesis stereopsis is an adaptation that enables animals to better visually locate and seize prey without moving (i.e. without motion parallax; (Cartmill, 1974; Cartmill, 1992)). Similarly, the camouflage-breaking hypothesis of stereopsis argues that stereopsis enables animals to detect camouflaged prey (Julesz, 1971).The visual predation theory was later elaborated to concern especially nocturnal animals as a large binocular visual field meant that the animals could better detect objects in dark environments (Pettigrew, 1986). However, the visual predation theory is undermined by the fact that some excellent predators do not have frontal eyes, and early nocturnal

primates may have detected their prey (mainly crawling insects) using olfaction or hearing (Sussman, 1991; Sussman, 2013; Crompton, 1995). Moreover, these smaller primates were largely omnivorous, eating lots of plant material in addition to insects, which lead Sussman ((1991), p. 219) to argue "that the explanation [for orbital convergence] will be found in adaptations providing the fine discrimination needed to exploit the small food items available on the newly diversifying flowering plants". Establishing the possible functional advantage provided by visual disparity is crucial to understand what factors may have driven the evolution of frontal eyes.

Random dot stereograms illustrate how stereopsis may enhance perception and enable the detection of otherwise hidden objects (Julesz, 1971), but they represent extremely simplified cases: all other depth cues (than binocular disparity) have been removed, and the object is placed in front of a flat background surface. Visual object recognition or masking studies have shown that binocular disparity can improve the accuracy of recognizing visual objects (Edelman & Blüthoff, 1992; Caziot & Backus, 2015; Lee & Saunders, 2011; Wardle, et al., 2010), but also in these studies the target objects were displayed in front of a flat background. In real life, the objects to be detected are embedded in scenes that include various depth cues (shadows, occlusion, object size etc.), and the scenes are often cluttered with other objects that occupy different depths. Hence, it is important to use stimuli that more closely resemble natural environments. There is anecdotal evidence that stereopsis helps observers to detect objects in realistic scenes as aerial stereographic photographs (often with greatly exaggerated visual disparity) have been used to detect camouflaged objects since the first world war (Saint-Amour, 2003). McKee et al. (1997) has shown that visual disparity aids in detecting a moving target dot that is flanked by randomly moving distractor dots that occupy a different plane in depth, but crucially when the moving target is presented in the midst of randomly moving distractors, disparity provides negligible help (McKee, et al., 1997). Valsecchi et al. (2013) reported that binocular disparity did not improve the accuracy of recognizing artificial or naturalistic scenes when monocular cues were available to provide information about the structure of the scene. However, McKee and Taylor (McKee & Taylor, 2010) showed that when real stimuli (instead of computer generated stimuli) were used, binocular depth discrimination thresholds of two adjacent objects were clearly superior to monocular thresholds.

To study if binocular disparity improves figure-ground segmentation in naturalistic environments, we asked participants to rapidly detect and enumerate individuals presented in an immersive virtual environment. Enumeration reveals well-documented limits in visual capacity: when 1-3 items are presented, enumeration is fast, accurate and relatively independent of number – a process termed *subitizing*. Larger sets decrease enumeration accuracy and increase response times (Trick & Pylyshyn, 1994; Railo, et al., 2016). The number of items that can be "subitized" is considered to mark the capacity of items that can be individuated and differentiated from the background and each other in a single glance (Jevons, 1871; Trick & Pylyshyn, 1994; Railo, et al., 2008; Piazza, et al., 2011). If binocular disparity increases the capacity of perception, enumeration should become faster and more accurate when compared to the condition where visual disparity has been removed (the same image is presented to both eyes). Figure 1 presents an overview of the present experiment and the hypotheses. One can hypothesize that the information added by binocular disparity enhances performance across different set sizes (Fig. 1B, left), for example, because disparity enables more efficient figure-ground segregation because the items can be resolved in depth when disparity is available. The influence of disparity could be limited to smaller set sizes, i.e. those items that the participants can efficiently process (Fig. 1B, middle). Third, it could be assumed that binocular disparity allows participants to augment the capacity of perception by enabling them to subitize larger set sizes (Fig. 1B, right). Although Fig. 1B only presents hypothesis concerning enumeration accuracy, similar hypotheses can be made also about reaction times. Response accuracy and speed may also reflect different aspects of visual capacity limits: Whereas response accuracy reflects the fidelity of the participant's response, reaction times may be more sensitive to the speed the participant is processing the items.

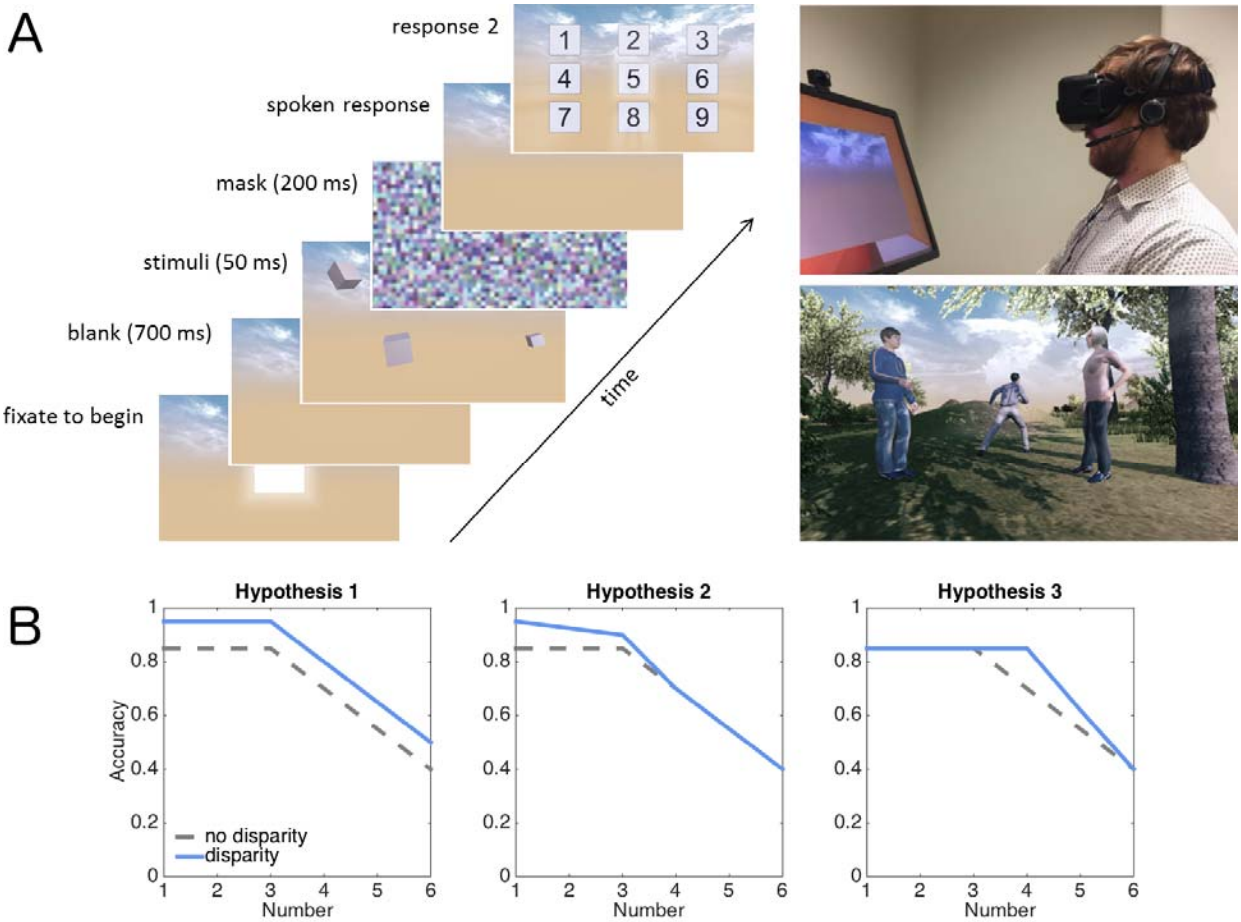

Figure 1. A) Task and stimuli. Schematic presentation of a single trial with cube stimuli (left; Experiment 2), the experimental set up (top right), and an example of a scene stimulus (Experiment 1). In the experiments, the stimuli (1-6 objects) were always presented binocularly and disparity was manipulated. In one condition the participants were asked to enumerate the individuals in the scene and in the other condition the participants rated the strength of their experience of depth. B) Hypotheses concerning enumeration accuracy.

A further interesting open question concerns to what degree the possible improvement in objective visual functioning (e.g. lower detection thresholds) due to binocular disparity is mirrored by changes in subjective, consciously experienced perception of depth. While many classical studies use the term stereopsis to refer both to the subjective (qualitative) experience of immersive depth and the objective (quantitative) coding of depth, Vishwanath (2014) has argued that the two senses of the term stereopsis should be distinguished. Consequently, the effects of binocular disparity can be measured in two ways. First, one can examine how binocular disparity influences subjective

experience of depth, for instance, by asking participants to rate the strength of their subjective experience of depth (such rating scales are commonly used in studies of conscious vision, see e.g. (Sandberg, et al., 2010; Seth, et al., 2008)). Second, one can measure how binocular disparity affects objective, visually-guided behavior (e.g. recognition accuracy and reaction times). Studies suggest that these two senses of the term stereopsis can be dissociated: Subjective sensation of depth can emerge without binocular disparity (Vishwanath & Hibbard, 2013), and disparity can influence behavior in the absence of subjective sensation of depth (Masson, et al., 1997; Cumming & Parker, 1997). Concerning the present study, we asked whether binocular disparity has similar effects on the subjective and objective sides of stereopsis. Or can binocular disparity increase the participants' subjective experience of depth without affecting their objective visual performance, or vice versa? Typically studies only measure objective visual performance and assume that binocular disparity also enhances the subjectively experienced impression of depth. However, given that the brain can construct a strong subjective sense of depth based on monocular depth cues (Vishwanath & Hibbard, 2013), one should not simply assume that subjective sense of depth is straightforwardly determined by binocular disparity.

The present study employed very brief presentation times (50 msec) combined with backward masking to focus on the contribution of information contained in visual disparity per se. Such brief stimulus durations minimize the contributions from motion parallax (if the person moves his head), but also limits the strength of subjective perception of depth (Westheimer, 1994). However, if visual disparity contributes to the speed and accuracy of segmenting objects from the background, the effect should be observed even with short presentation times that may be insufficient to produce a strong subjective sensation of depth. The amount of visual disparity was also restricted in the present study to minimize the physical difference between the disparity vs. no-disparity conditions (see Methods). This meant that from the participant's point-of-view the disparity vs. no-disparity conditions looked very similar (none of the participants reported spontaneously perceiving a change in disparity during the experiment).

Method

Participants

Twenty-nine participants took part in Experiment 1 (age range 18-34, 23 females), and 31 participants to Experiment 2 (age range 19-35, 26 females). All participants reported having normal or corrected-to-normal vision, and reported that they did not have deficits in stereo vision (e.g. strabismus). Participants gave informed consent to participate the study, and the study was conducted in accordance with the Declaration of Helsinki.

Stimuli and procedure

Participants were presented with stereoscopic and non-stereoscopic stimuli in immersive virtual reality using the Oculus Rift DK2 headset (resolution 960 × 1080 per eye, 75 Hz refresh rate, 100° field of view). Immersive virtual reality refers to the fact that the participant feels being present in the virtual world. The world is immersive in the sense that it surrounds the participant and by turning her head the participant can look around in the virtual world. The Oculus Rift was running on a computer (3.50 GHz, 16 GB RAM) with a NVIDIA GeForce GTX 980 graphics card. The camera that tracks the position and orientation of the headset was placed roughly on the table roughly 1 meter from the participant.

In Experiment 1 the stimuli were natural scenes that contained 1–6 humans (see Fig. 1). The humans were generated using an open source Make Human software (www.makehuman.org). They were distributed around the scene, they never overlapped with each other, and at least their head and torso were clearly visible. The scenes, which were generated using various “prefabs” of the Unity software, contained trees, bushes or rocks, but never any artificial, man-made objects. The humans were positioned directly in front of the camera. In Experiment 2 the stimuli were randomly oriented cubes presented at pseudo-random positions in front of the participant on sky background (same as in Experiment 1). The color of the cubes was constant within a scene, but varied across scenes.

The experiment and the stimuli were made using the game engine Unity (5.4.0f3). The images presented to the left and right eye screens in the Oculus Rift were determined by two cameras placed in the virtual world constructed by Unity. We used the default settings: the distance between the left

and right virtual cameras was 64 mm, and they were parallel to each other, that is, the cameras were not angled inward like eyes typically are. While this means that the size of visual disparity is reduced (which in turn is likely to reduce the size of the effect disparity has on enumeration performance), the advantage is that the scenes with and without visual disparity look very similar (the manipulation of disparity was not spontaneously noticed by the participants). This is important as strong variation in disparity could draw the participants' attention and influence behavior indirectly. Furthermore, stronger disparity could produce nausea in some participants. In the "No disparity" condition the image of the left virtual camera was presented to both left and right eyes of the participant.

As shown in Figure 1, each trial began with the presentation of a light gray square positioned directly in front of the participants. When the participant oriented towards the square for 1 second it turned to a darker shade and triggered the presentation of the scene 700 msec later. Each scene was presented for 50 msec, and it was followed by a flat (i.e. 2D) pixelated mask image for 200 msec. In the enumeration task the participant was asked to say out loud the number of humans presented as quickly as possible. The participants wore Sennheiser PC-8 headphones with a microphone. Reaction times relative the onset of the scene were recorded. After the vocal response numbers 1–9 were presented on a 3 x 3 grid of squares in the virtual reality. The participant selected the number she had spoken by orienting towards it in the virtual reality. The answer was selected when the participant kept orienting towards the number for at least 1 second. In the subjective experience of depth task the participant was asked to rate the strength of the subjective experience of depth and three-dimensionality on a scale from one to three. The highest alternative denoted a strong sense three-dimensionality, whereas the lowest alternative meant that the participant experienced a relatively flat, 2D world with very little sensation of depth. The middle alternative referred to medium strength experience of depth. We also emphasized to the participant that this task was not related to the apparent quality of the graphics, nor was there a correct answer to the question. The participant was told that she did not have to make a speeded response. The answer was given first by speaking to a microphone and then selecting the corresponding response by orienting to it in the virtual reality.

Each participant performed two blocks of 60 trials of the enumeration and two blocks of 60 trials of the depth rating task. The order of the tasks was counterbalanced across participants. Each

participant saw each stimulus once, and across participants each stimulus appeared equally often in the enumeration and in the depth rating tasks. Ten practice trials were completed before each task.

Statistical analysis

Data were analyzed in R statistical software (version 3.1.2; (R Development Core Team , 2014)). Single-trial depth rating and reaction time data was analyzed using linear mixed-effects models (Baayen, et al., 2008), and accuracy data with logit mixed-effects models (Jaeger, 2008) with binomial probability distributions (using the lme4 package; (Bates, et al., 2014)). Mixed-effects models enable taking into account individual differences (e.g., in the speed of responding) in the model, in addition to group-level effects. In the mixed-effect models, the number of items (1–6), binocular disparity (disparity vs. no disparity), and their interaction were added as fixed-effects factors. In addition, the reaction time of a previous trial was included to reaction time models, and the depth rating of a previous trial was included to the depth rating analyzes. Both of these continuous variables were centered. Participants were defined as random effects so that each participant's data were fitted with individual intercepts and slopes (for both disparity and number conditions). Moreover, because enumeration data is known to contain a breakpoint around number 2-4, the data was fitted with two-segment linear models in addition to standard one segment models (for formal description, see (Railo, et al., 2016)). We fitted three alternative two-segment models where the length of the first segment ranged numbers 1–2, 1–3, or 1–4. These were compared against the one segment model with a corresponding random-effect structure. Model fits were evaluated using the Akaike Information Criterion, which weighs the goodness of fit by the complexity of the model. Different models were compared using likelihood-ratio tests. The best fitting model was pruned by removing nonsignificant regressors ($t < |2|$). The pruned models are visualized in the figures. We do not report p values because there is no standardized way to calculate the degrees of freedom and the corresponding p values in mixed-effects models (Baayen, et al., 2008). The statistical significance of an effect can be checked from the confidence intervals (CI): whenever the value zero is included in the CI, the effect is not considered statistically significant. Error bars in all the figures are 95% CIs calculated from 1000 bootstrap samples.

Datasets and the analysis scripts can be downloaded at the Open Science Framework (https://osf.io/9b3h8/).

Results

Experiment 1: Natural scenes

The results of Experiment 1 where the participants viewed natural scenes are presented in Figure 2. At group level, subjective depth ratings decreased .20 units on average when binocular disparity was removed (β = -0.20, CI= [-0.25, -0.15], t = 8.5). Moreover, depth ratings increased with number (β = 0.065, CI = [-0.043, 0.087], t = 5.9). The interaction between number and disparity condition was not statistically significant, and was thus left out of the model. Finally, the depth rating of previous trial predicted ratings of depth (β = -0.089, CI = [0.058, 0.12], t = 5.6).

Enumeration accuracy was best fitted with a two segment model (χ^2 = 12.1, df = 4, p = .016): Accuracy remained high for the first three items (subitizing), after which it declined rapidly. Accuracy was slightly lower without binocular disparity during the first segment (OR = 0.66, CI = [0.49, 0.90], Z = -2.61). Accuracy also slightly decreased as a function of number for the first three numbers (OR = 0.45, CI = [0.38, 0.55], Z = -8.04). In the second segment, enumeration accuracy declined more steeply as a function of number (OR = 0.54, CI = [0.40, 0.73], Z = -3.9). An interaction between disparity condition and number (4-6) in the second segment indicated that enumeration accuracy was initially higher in the second segment during binocular disparity (OR = 1.44, CI = [1.13, 1.86], Z = 3.00). This result is important as it shows that during binocular disparity enumeration remains accurate longer than without disparity. In addition to modelled results, this result is also clearly seen in the observed data: Bonferroni corrected pairwise t tests between the disparity conditions at each number confirmed that a statistically significant difference was only present at number four (difference 8.5 percentage units, p = .002; other comparisons: uncorrected p values ≥ .37).

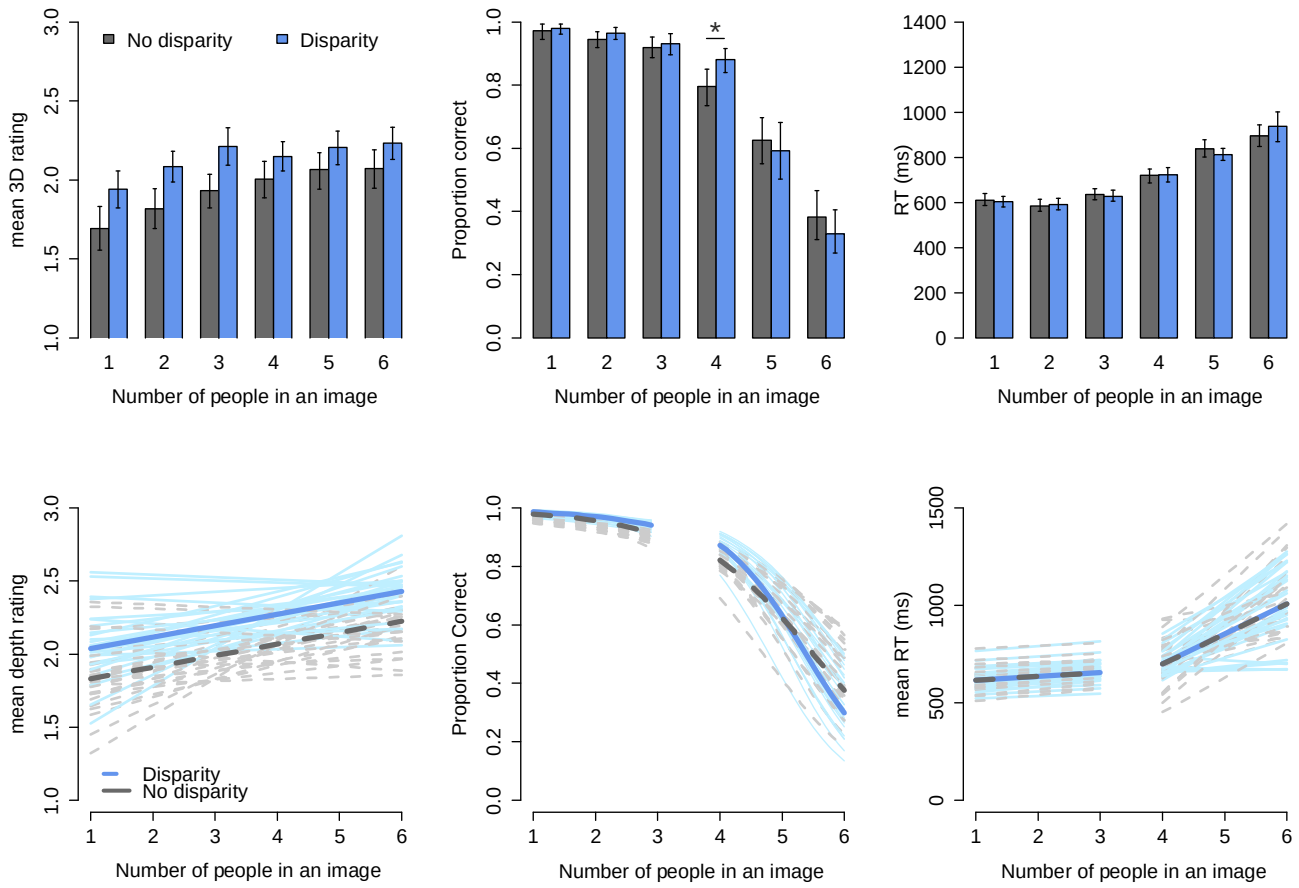

Figure 2. Results of Experiment 1: Observed data (upper panels) and modelled data (lower panels) when participants viewed natural scenes. Error bars show 95% CIs. In the lower panels, the bold lines depict the group-level results and the thinner lines in the background show individual participants' results (i.e. random-effect variation).

Reaction times were also best modelled by a two-segment model rather than a corresponding one segment linear model (χ^2 = 354.1, df = 37, p < .0001). Reaction times slightly increased for the first three numbers (β = 13.14 ms, CI = [6.46, 19.81], t = 3.86; intercept: β = 616.03 ms, CI = [596.01, 636.06], t = 60.30). For the second segment enumeration times increased more steeply (β = 93.46 ms, CI = [73.299, 105.62], t = 10.85). Thus, whereas binocular disparity increased enumeration accuracy, it did not influence the speed of correct responses.

Experiment 2: Cube stimuli

Experiments 1 suggests that when viewing natural scenes binocular disparity may increase visual capacity without influencing the speed of visual processing. Experiment 2 was designed to test whether these effects are larger when the stimuli contain very few other depth cues than binocular disparity. In Experiment 2 we employed geometric cubes presented against a distant sky background as the stimuli to be enumerated (see Figure 1). Here other depth cues (than binocular disparity) are minimized whereas the natural scene stimuli employed in Experiment 1 contains various depth cues (e.g. familiar objects in predictable environments, occlusion, shading). We expected this manipulation to enhance the effect binocular disparity has on enumeration performance.

The results are shown in Figure 3. Depth ratings were best fitted by a two-segment model where the first segment included numbers 1-3 ($\chi^2 = 27.3$, df = 4, $p < .0001$). Surprisingly, participants' subjective experience of depth was not modulated by binocular disparity. For numbers 1-3, mean depth rating was 1.90 ($\beta = 1.90$, CI = [1.80,2.00], t = -37.6), and for numbers 4-6 depth ratings increased by 0.24 ($\beta = 0.24$, CI = [0.18, 0.31], t = -7.3).

The accuracy of enumeration in Experiment 2 was very similar to the results obtained in Experiment 1 with day time scenes as stimuli. Enumeration accuracy was best fit by a two-segment model ($\chi^2 = 15.1$, df = 1, $p < .0001$). In the first segment, enumeration accuracy decreased a function of number in the first segment (OR = 0.24, CI = [0.18, 0.39], Z = -9.6), and enumeration accuracy was lower when binocular disparity was eliminated (OR = 0.61, CI = [0.43, 0.87], Z = -2.7). Crucially, the second segment also revealed an interaction between number and disparity condition (OR = 1.42, CI = [1.09, 1.84], Z = -2.6): Enumeration remains accurate longer with binocular disparity. Bonferroni corrected t tests at each number comparing the two disparity conditions verified that enumeration accuracy was reduced without binocular disparity (difference 14.3 percentage units; $p = .0048$) only at number four (other uncorrected p values $\geq .48$). All in all, these results with cube stimuli replicate the results obtained with day time scenes concerning enumeration accuracy.

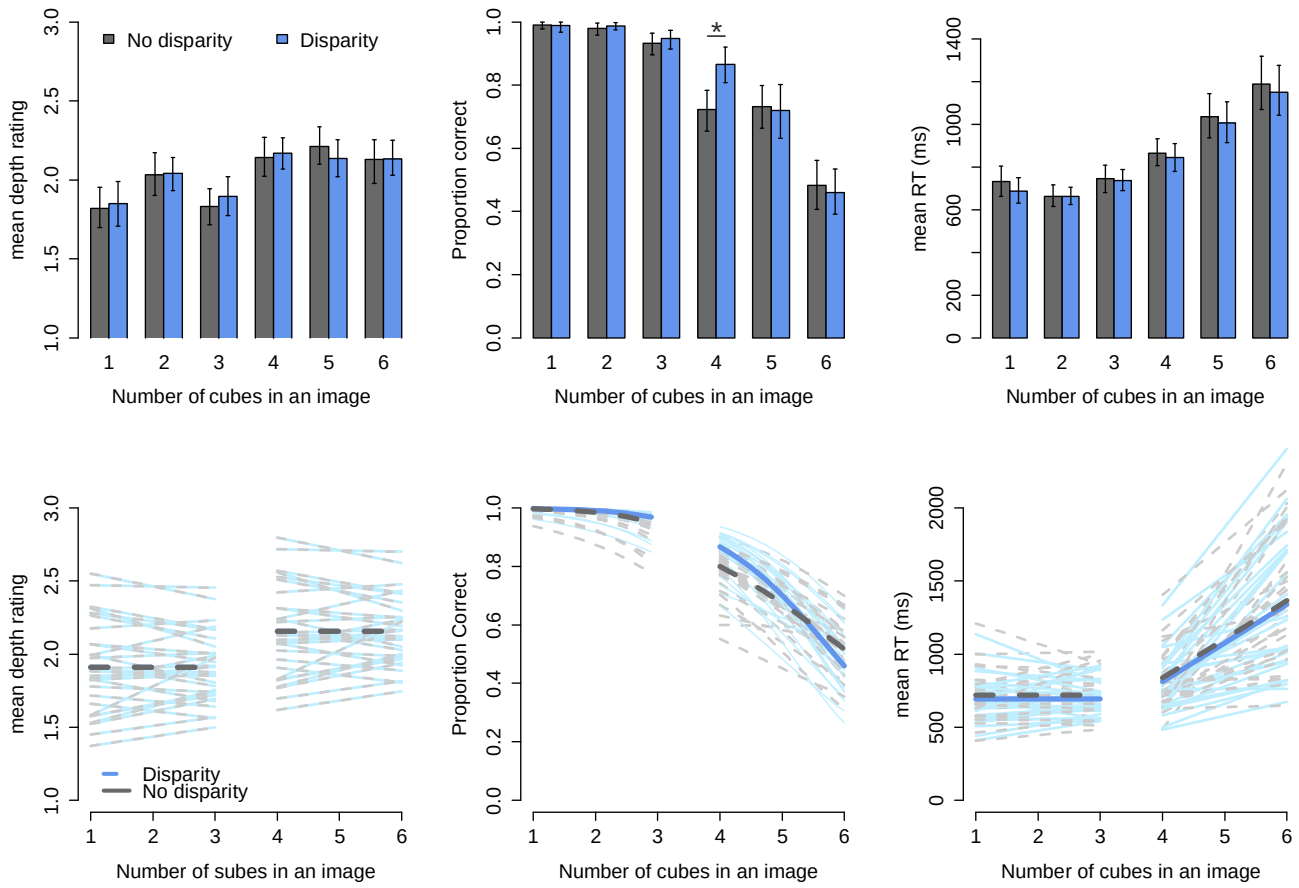

Figure 3. Results of Experiment 2: Observed data (upper panels) and modelled data (lower panels) when participants viewed geometric shapes. Error bars show 95% CIs. In the lower panels, the bold lines depict the group-level results and the thinner lines in the background show individual participants' results.

In Experiment 2 reaction times were delayed by a constant of 27 ms when binocular disparity was removed (see Fig 3; β = 27.04 ms, CI = [8.08, 45.99], t = 2.7). That is, with cubes as stimuli, enumeration was *faster and more accurate* with disparity than without it. Otherwise, the reaction times followed a similar two-segment pattern (χ^2 = 310.2, p < .0001) as previously. In the first segment reaction times remained constant (β = 693.2 ms, CI = [647.66, 738.81], t = 29.8). During the second segment reaction times increased as a function of number (β = 175.8 ms, CI = [130.71, 220.82], t = 7.6).

Discussion

Different theories concerning the evolution of frontal eyes have in common the assumption that binocular disparity is an adaptation that enables the animals to better discriminate objects from the background (Crompton, 1995). The present study examined if binocular disparity aids the enumeration of individual objects in cluttered, naturalistic scenes in immersive virtual reality. Enumeration paradigm was employed because we were interested if the influence of binocular disparity on recognition or figure-ground-segregation showed set-size dependent effects. Previous studies have examined how binocular disparity influences the processing of singular items displayed against a flat background. Furthermore, we examined the effects of binocular disparity on objective visual performance (enumeration accuracy and reaction times) and subjectively reported experience of depth. Our results show that even very slight binocular disparity, one that is not noticed by the participant and does necessarily produce a concurrent enhancement of consciously perceived depth, can augment the capacity of visual perception in that more visual objects can be individuated at a glance. However, notably our results suggest that the effect of binocular disparity on visual enumeration accuracy and reaction times is very limited; it occurs at the "limit" of perceptual capacity. Hence, consistent with Valsecchi et al. (2013), our results suggest that disparity does not produce a strong general benefit for visual object individuation. However, arguably the capacity added by binocular disparity may have helped animals to better individuate objects (e.g. fruit, prey, predators) that are embedded among other objects (e.g. plants and branches) in complex visual environments (Sussman, 2013). Our results also raise the question if individuals with impaired binocular vision have a reduced visual perceptual capacity.

With binocular disparity, the accuracy of enumeration was increased when the scene contained four target individuals, indicating that binocular disparity enhanced the accuracy of perception right at the limit of the subitizing capacity. Because disparity only affected enumeration accuracy when the scene contained four items, one is led to ask if our observation is reliable. Given that virtually identical results were obtained in both experiments using different stimuli, we conclude that the observed result is reliable. Moreover, both experiments had a relatively large sample size and the effect observed was very systematic across individual participants (Bonferroni corrected p values < .005). At number four, disparity increased enumeration accuracy on average by 8.5 and 14.2 percentage units

in Experiments 1 and 2, respectively. The size of this results is similar to the improvement observed when experienced video-game players are compared to non-gamers – as in the present study, the observed increase in enumeration accuracy was about 10 percentage units, and began at number four (Green & Bavelier, 2003; Green & Bavelier, 2006). Note that although the modelled results suggest that binocular disparity may decrease enumeration accuracy at high numbers (with six items), this effect is likely to represent a byproduct of the (nonlinear) logit model used to model the accuracy data. Importantly, similar effect is not observed in the empirical data. The mixed-effects models also suggested that disparity increased enumeration accuracy also in the subitizing range, but the size of this effect was minute.

When geometric cube stimuli were used, disparity also sped-up reaction times by 27 ms on average. While the absolute size of this effect on reaction times is relatively small, it may represent a significant advantage in situations where fast visually-guided behavior is important (e.g. see Julesz, 1971, for similar argument). However, given that this effect was only observed when cube stimuli were used, the present results suggest that this benefit may not generalize to real-world situations. The reader should also note that logically one could have also expected to observe longer reaction times under binocular disparity, as integrating the two different viewpoints of the eyes is a complicated task that can be assumed to require additional processing time. Reaction times were enhanced only in Experiment 2 possibly because other depth cues were minimized with the use of geometric shapes.

We propose that superior accuracy in enumerating four objects with disparity is produced by the fact that disparity enabled the participants to better individuate the objects from other objects and the background. Franconeri et al. (2013) have suggested that capacity limits in perception result from competition between objects in representational maps: adjacent items inhibit each other and compete for limited “cortical real estate” (Cohen, et al., 2015). Binocular disparity may counteract the competition between different objects within representational maps, enabling the visual system to more efficiently process the objects. This advantage is small when visual discrimination is relatively easy (i.e. within the subitizing range), but the advantage produced by disparity becomes apparent in situations where the number of items is just above the limits of visual processing capacity (i.e. four items). This effect could arise at earliest in V1, which is the first stage of visual processing that

contains disparity selective neurons, although higher cortical areas may also contribute to the effect (Barlow, et al., 1967; Cumming & DeAngelis, 2001). Stereoscopic perception constrains the allocation of selective visual attention (He & Nakayama, 1995), and as subitizing is known to depend on attentional resources (Railo, et al., 2008), the better enumeration performance under binocular disparity could be due to easier deployment of attention to the objects to be enumerated. Why does the beneficial effect of disparity only show up at number four, and not at higher numbers? A possible explanation is that one the number of items is high enough, the beneficial effect of disparity on representational maps does not yield a strong enough influence. That is, once set size is high, the competition in representational maps (i.e. inhibition between objects) is so high that disparity cannot reverse it.

Interestingly, our results show that the advantage produced by binocular disparity does not require an enhanced conscious perception of depth, suggesting that the mechanisms that produce the functional advantage are somewhat independent from the processes that produce the conscious experience of depth often associated with disparity. The dissociation between enhanced subjective perception of depth and the functional advantage provided by binocular disparity is in concert with recent theoretical proposals that emphasize the difference between functional and subjective forms of stereopsis (Vishwanath, 2014), and with the fact that binocular disparity is not necessary for subjective sensation of depth (Vishwanath & Hibbard, 2013). Moreover, binocular disparities can elicit vergence eye movements in the absence of subjective perception of depth (Masson, et al., 1997). While vergence eye movements and the behavioral effects observed in the present study could be mediated by disparity-selective neurons in V1, these neurons are unlikely to account for subjective perception of depth (Cumming & Parker, 1997). We suggest that while the functional advantage produced by binocular disparity is likely produced by following rapid computations (possibly in V1; (Masson, et al., 1997; Cumming & Parker, 1997)), the encoding of subjective experience of depth requires longer processing times and integration of information across different cortical areas (Koch, et al., 2016). Further studies are required to pinpoint the neural mechanisms involved in the two processes. For instance, whereas the functional advantage provided by disparity could also be mediated predominantly by fast acting dorsal visual stream/magnocellular neurons, but the ventral

stream/parvocellular neurons may contribute more to the subjective experience of depth (Parker, 2007).

The present study also has important limitations. First, further studies are required to replicate the present findings with other stimuli. Future experiments should aim to control and vary stimulus configurations systematically to examine how various stimulus features (density, field-of-view etc.) affect the results. It is unclear whether different results would be observed if the scenes were nocturnal (Pettigrew, 1986), which may decrease the visual systems ability to use monocular depth cues. Moreover, while the scenes used in the present study included unrelated objects (e.g. trees, rocks), it remains open disparity affects visual performance in more cluttered contexts (e.g. in the midst of tree branches). Also, it remains to be tested whether the visual system is more sensitive to binocular depth cues when, for example, either threatening or otherwise "evolutionarily relevant" (i.e. snakes, see (Isbell, 2009)) are present. Second, while we aimed to use naturalistic stimuli in Experiment 1, and employed virtual reality technology, the stimuli are nevertheless artificial. Use of real stimuli (real individuals in real scenes) could produce different results (McKee & Taylor, 2010). For instance, in real life, vergence eye movements provide information about distance, whereas with virtual reality headsets the eyes are always focused on the screen surface. Third, future studies should aim to systematically vary object locations, as objects which are closer to the observer (e.g. at reaching distance) are likely to get more benefit from binocular disparity. Similarly, there is reason to assume that the size of the advantage produced by binocular disparity would be larger if binocular disparity would be increased by aligning the virtual cameras (i.e. eyes) inward (producing larger disparity), and if the scene would be presented for a longer duration (enabling longer accumulation of visual information). We purposely limited the amount of disparity and presentation duration so that the participants would not notice the experimental manipulation (change in disparity). Strong variation in disparity could make the participant notice the manipulation, and draw their attention to it, and thereby influence behavior indirectly. None of the participants in the present experiments reported noticing the disparity manipulation.

In conclusion, the present results suggest that binocular disparity may enhance the capacity of visual perception by increasing the number of objects that can be individuated from the background. This effect can arise without concurrent enhancement in consciously experienced depth.

Acknowledgements

The research was supported by the Finnish Cultural Foundation, and the Academy of Finland (H.R., grant #308533). We thank Iiris Ollikainen for help with data collection.

Author contributions

H.R. designed the research, analyzed the data, and wrote the article. A.P. programmed the experiments, J.S. and S.K. collected the data.

Competing interests

No competing interests declared

References

Baayen, R. H., Davidson, D. J. & Bates, D. M., 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language,* Volume 59, pp. 390-412.

Barlow, H. B., Blakemore, C. & Pettigrew, J. D., 1967. The neural mechanism of binocular depth discrimination. *The Journal of Physiology,* 193(2), pp. 327-342.

Barry, S. R., 2009. *Fixing My Gaze.* New York(NY): Basic Books.

Barton, R. A., 2004. Binocularity and brain evolution in primates. *Proceedings of the National Academy of Sciences of the USA,* 101(27), pp. 10113-10115.

Bates, D., Maechler, M., Bolker, B. M. & Walker, S., 2014. lme4: Linear mixed-effects models using Eigen and S4. *R package version 1.1-7, URL: http://CRAN.R-project.org/package=lme4.*

Cartmill, M., 1974. Rethinking primate origins. *Science,* Volume 184, pp. 436-443.

Cartmill, M., 1992. New views on primate origins. *Evolutionary Anthropology,* Volume 1, pp. 105-111.

Caziot, B. & Backus, B. T., 2015. Stereoscopic offset makes objects easier to recognitze. *PlosOne,* p. doi.org/10.1371/journal.pone.0129101 .

Cohen, M. A., Rhee, J. Y. & Alvarez, G. A., 2015. Limits on perceptual encoding can be predicted from known receptive field properties of human visual cortex. *Journal of Experimental Psychology: Human Perception and Performance,* 42(1), pp. 67-77.

Collins, E. T., 1921. Changes in the visual organs correlated with the adoption of arboreal life and with the assumption of the erect posture. *Transactions of the Opthalamological Society of the UK,* Volume 41, pp. 10-90.

Crompton, R. H., 1995. "Visual predation," habitat structure, and the ancestral primate niche. In: L. Alterman, G. A. Doyle & M. K. Izard, eds. *Creatures of the Dark: The Nocturnal Prosimians.* New York(New York): Plenum Press, pp. 11-30.

Cumming, B. G. & DeAngelis, G. C., 2001. The physiology of stereopsis. *Annual Review of Neuroscience,* Volume 24, pp. 203-238.

Cumming, B. G. & Parker, A. J., 1997. Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature,* Volume 389, pp. 280-283.

Dal Mutto, C., Zanuttigh, P., Cortelazzo, G. M. & Mattoccia, S., 2011. Scene segmentation assisted by stereo vision. *International Conference on 3D Imaging, Modeling, Processing, Visualization, and Transaction,* pp. 57-64.

Edelman, S. & Blüthoff, H. H., 1992. Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Research,* 32(12), pp. 2385-2400.

Fielder, A. R. & Moseley, M. J., 1996. Does stereopsis matter in humans. *Eye,* Volume 10, pp. 233-238.

Franconeri, S. L., Alvarez, G. A. & Cavanagh, P., 2013. Flexible cognitive resources: competitive content maps for attention and memory. *Trends in Cognitive Sciences,* 17(3), pp. 134-141.

Green, C. S. & Bavelier, D., 2003. Action video game modifies visual selective attention. *Nature,* Volume 423, pp. 534-537.

Green, C. S. & Bavelier, D., 2006. Enumeration versus multiple object tracking: The case of action video game players. *Cognition,* Volume 101, pp. 217-245.

Heesy, C. P., 2009. Seeing in stereo: The ecology and evolution of primate binocular vision and stereopsis. *Evolutionary Anthropology,* Volume 18, pp. 21-35.

He, Z. & Nakayama, K., 1995. Visual attention to surfaces in three-dimensional space. *Proceedings of the National Academy of Sciences of USA,* Volume 92, pp. 11155-11159.

Isbell, L., 2009. *The Fruit, the Tree and the Serpent.* Cambridge(Massachusetts): Harvard University Press.

Jaeger, T. F., 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language,* 59(4), pp. 434-446.

Jevons, W. S., 1871. The power of numerical discrimination. *Nature,* Volume 3, pp. 281-282.

Jones, F., 1917. *Arboreal Man.* New York(NY): Hafner Publishing.

Julesz, B., 1971. *Foundations of Cyclopean Perception.* Cambridge, Massachusetts(London, England): MIT Press.

Koch, C., Massimini, M., Boly, M. & Tononi, G., 2016. Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience,* 17(5), pp. 307-321.

Lee, Y. L. & Saunders, J. A., 2011. Stereo improves 3D shape discrimination even when rich monocular shape cues are available. *Jopurnal of Vision,* 11(6), p. doi: 10.1167/11.9.6.

Masson, G. S., Busettini, C. & Miles, F. A., 1997. Vergence eye movements in response to binocular disparity without depth perception. *Nature,* Volume 389, pp. 283-286.

McKee, S. P. & Taylor, D. G., 2010. The precision of binocular and monocular depth judgements in natural settings. *Journal of Vision,* 10(10), p. doi:10.1167/10.10.5.

McKee, S. P. et al., 1997. Is stereopsis effective in breaking camouflage for moving targets. *Vision Research,* 37(15), pp. 2049-2055.

Melmoth, D. R., Finlay, A. L., Morgan, M. J. & Grant, S., 2009. Grasping deficits and adaptations in afults with stereo vision losses. *Investigative Ophalamology & Visual Science,* 50(8), pp. 3711-3720.

O'Connor, A. R. et al., 2010. The functional significance of stereopsis. *Investigative Opthalamology & Visual Science,* 51(4), pp. 2019-2023.

Parker, A. J., 2007. Binocular depth perception and the cerebral cortex. *Nature Reviews,* Volume 8, pp. 379-391.

Pettigrew, J. D., 1986. The evolution of binocular vision. In: J. D. Pettigrew, K. J. Sanderson & W. R. Levick, eds. *Visual Neuroscience.* London: Cambridge University Press, pp. 208-222.

Piazza, M., Fumarola, A., Chinello, A. & Melcher, D., 2011. Subitizing reflects visuo-spatial object individuation capacity. *Cognition,* Volume 121, pp. 147-153.

R Development Core Team , 2014. *R: A language and environment for statistical computing.* Vienna(Austria): R Foundation for Statistical Computing. Available from http://www.R-project.org/.

Railo, H. et al., 2016. Rapid and accurate processing of multiple objects in briefly presented real-world images. *Journal of Vision,* Volume 16, p. doi:10.1167/16.3.8..

Railo, H., Koivisto, M., Revonsuo, A. & Hannula, M. M., 2008. The role of attention in subitizing. *Cognition,* Volume 107, pp. 82-104.

Ross, C. F. & Martin, R. D., 2006. The role of vision in the origin and evolution of primates. In: J. H. Kaas, ed. *Evolution of Neurvous Systems.* Oxford: Academic Press, pp. 59-78.

Saint-Amour, P., 2003. Modernist Reconnaissance. *Modernism/modernity,* 10(2), pp. 349-380.

Sandberg, K., Timmermans, B., Overgaard, M. & Cleeremens, A., 2010. Measuring consciousness: Is one measure better than the other?. *Consciousness and Cognition,* 19(4), pp. 1069-1078.

Seth, A. K. et al., 2008. Measuring consciousness: Relating behavioural and neurophysiological approaches. *Trend in Cognitive Sciences,* 12(8), pp. 314-321.

Smith, G. E., 1924. *The Evolution of Man.* London: Oxford University Press.

Sussman, R. W., 1991. Primate origins and the evolution of angiosperms. *American Journal of Primatology,* Volume 23, pp. 209-223.

Sussman, R. W., 2013. Rethinking primate origins again. *American Journal of Primatology,* Volume 75, pp. 95-106.

Trick, L. M. & Pylyshyn, Z. W., 1994. Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review,* 101(1), pp. 80-102.

Valsecchi, M., Caziot, B. T., Backus, B. T. & Gegenfurtner, K. R., 2013. The role of binocular disparity in rapid scene and pattern recognition. *i-Perception,* Volume 4, pp. 122-136.

Wardle, S. G., Cass, J., Brooks, K. R. & Alais, D., 2010. Breaking camouflage: Binocular disparity reduces contrast masking in natural images. *Journal of Vision,* 10(14), pp. 1-12.

Westheimer, G., 1994. The Ferrier Lecture, 1992. Seeing depth with two eyes: stereopsis. *Proceedings of the Royal Society B: Biological Sciences,* Volume 257, pp. 205-214.

Vishwanath, D., 2014. Toward a new theory of stereopsis. *Psychological Review,* Volume 121, pp. 151-178.

Vishwanath, D. & Hibbard, P. B., 2013. Seeing in 3-d with just one eye: stereopsis without binocular vision. *Psychological Science,* 24(9), pp. 1673-1685.