Vision Research 125 (2016) 64–75

Contents lists available at ScienceDirect

Vision Research

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

Depth magnitude from stereopsis: Assessment techniques and the role of experience

Brittney Hartle *, Laurie M. Wilcox

Centre for Vision Research, Lassonde Building, York University, 4700 Keele Street, Toronto, ON, Canada

ARTICLE INFO

Article history:
Received 9 October 2015
Received in revised form 12 February 2016
Accepted 16 May 2016
Available online 6 July 2016

Keywords:
Stereopsis
Depth
Depth magnitude
Quantitative depth

ABSTRACT

Investigations of the relationship between binocular disparity and suprathreshold depth magnitude percepts have used a variety of tasks, stimuli, and methods. Collectively, the results confirm that depth percepts increase with increasing disparity, but there are large differences in how well the estimates correspond to geometric predictions. To evaluate the source of these differences, we assessed depth magnitude percepts for simple stereoscopic stimuli, using both intra- and cross-modal estimation methods, and a large range of test disparities for both experienced and inexperienced observers. Our results confirm that there is a proportional relationship between perceived depth and binocular disparity; this relationship is not impacted by the measurement method. However, observers with minimal prior experience showed strong systematic biases in depth estimation, which resulted in large overestimates at small disparities and substantial underestimates at large disparities. By comparison, experienced observers' depth judgements were much closer to geometric predictions. In subsequent studies we show that unpracticed observers' depth estimates are improved by removing conflicting depth cues, and the observed biases are eliminated when they view physical targets. We conclude that differences in the depth magnitude estimates as a function of disparity in the existing literature are likely due to observers' experience with stereoscopic display systems in which binocular disparity is manipulated while other depth cues are held constant.

1. Introduction

The relationship between binocular disparity and the phenomenon of stereoscopic depth perception was firmly established early in the 19th century. Replications of Helmholtz's threshold discrimination study (using three lines) confirmed that, within Panum's fusional area, the best thresholds are as low as 2–5 arcsec (among others see Andersen & Weymouth, 1923; Helmholtz, 1925; Howard, 1919). Within this range of 'fusable' disparities, stereopsis has also been shown to support reliable and accurate depth magnitude judgements (Ogle, 1952, 1953), though depth percepts increase with increasing disparity for a range of diplopic disparities as well (Foley, Applebaum, & Richards, 1975; Ogle, 1953). While its precision has largely dominated stereoscopic research in the past 50 years, it is arguable that the suprathreshold properties of stereoscopic depth perception are just as relevant, if not more so, to natural tasks such as navigation, reaching, and grasping. However, as outlined by Foley et al. (1975) it is not possible to simply predict suprathreshold percepts from discrimination thresholds. Ogle (1953) and Foley et al. (1975) have assessed depth magnitude percepts over a wide range of disparities, making a special effort to control factors other than binocular disparity, that could influence observers' estimates. For instance, in his study of the 'precision and validity' of depth from large disparities Ogle (1953) eliminated factors such as relative size, blur and convergent eye movements, and manipulated eccentricity. In their experiments, Foley and Richards (1972) controlled these variables and manipulated exposure duration to assess the impact of vergence on suprathreshold depth estimates.

Taken together, the results of Ogle's, (1952, 1953) and of Foley and colleagues' experiments (Foley, 1968; Foley & Richards, 1972; Foley et al., 1975) show that when stimuli are positioned close to the fovea, depth percepts scale with increasing disparity over a large range of disparities. However, there have been a variety of patterns of bias reported by these authors. For instance, Foley and Richards (1972) assessed depth from relatively small disparities (as low as 10 arcmin) and their results show that in this range when eye movements are permitted, depth is slightly overestimated. Ogle (1953) tested disparities ranging from 12 to 80 arcmin, and his results show no such overestimation, though viewing time was restricted in his study. The minimum test disparity used by Foley et al. (1975) was close to 0.5 deg, and in both this and the work of Ogle (1953) depth estimates were lower than predicted from binocular viewing geometry; at very large, diplopic test disparities depth magnitude estimates no longer scaled propor-

* Corresponding author.
E-mail address: brit1317@yorku.ca (B. Hartle).

http://dx.doi.org/10.1016/j.visres.2016.05.006
0042-6989/

tionally with disparity. Between 0.5 and 2 deg the depth estimates reported by Foley et al. (1975) are substantially lower than predicted, by a factor of 4 in the crossed direction, where Ogle's estimates are only slightly below predicted levels. Foley et al. (1975) also note that there is a substantial reduction in perceived depth magnitude for uncrossed disparities, which they attribute to the nature of the virtual display, and to the manual pointing task used to assess perceived depth (see below).

As noted above, in these series of studies care was taken to eliminate or control factors that may have influenced depth magnitude estimates from disparity. An important consideration in all cases was the nature of the task used to quantify magnitude percepts. Depth estimation studies have used a variety of tasks, many of which have significant drawbacks. For instance, verbal reports of units (e.g. centimetres) have been shown to be highly sensitive to experimental context and response biases caused by experimental restrictions on the range of available responses (Poulton, 1968). In addition, verbal estimates are derived from an unspecified function of the depth from disparity estimate (i.e. output mapping problem). Moreover, unit estimation results from verbal estimates exhibit large interobserver variability that may be due to unit recall limitations rather than perceived depth per se (Foley et al., 1975).

While depth matching tasks have often been used to assess stereopsis, their results must be interpreted carefully because they do not quantify the perceived magnitude of a percept, they can only reflect that a given perceptual magnitude is equivalent to another (Foley et al., 1975; for review of these issues see Howard & Rogers, 2012). As an alternative to matching, Ogle (1952, 1953) used ratio-based judgements in which observers were asked to position an object at half of the depth between two targets, or to position an object in front of the fixation plane to represent the apparent distance of another stimulus positioned behind the fixation plane. These tasks require that observers estimate the amount of depth between a target and a reference plane. Foley et al. (1975) used a manual-pointing task in which observers were asked to point with an unseen finger at the position of a flashed target relative to the fixation plane. While this task seems more natural, as the authors allow, it may have introduced biases due to a tendency for observers to under-reach to large uncrossed disparities. Moreover, it is possible that observers may have been limited by their memory for the position of the very brief (40 ms) target flash. Another potentially important, but as yet unremarked difference between the work of Ogle and that of Foley and colleagues was their observers' prior experience with stereoscopic stimuli. Foley et al. (1975) noted that Ogle (1953) reported the results of only two observers (one of whom is the author), and they countered this by testing a larger set of individuals. However, they did not subsequently consider that differences between their data and those of Ogle might have been due to the characteristics of these observers, specifically their limited experience with such tasks.

Given these differences in stimuli, task, and range of test disparities it is difficult to compare the results of extant depth magnitude studies. In particular, while there is broad agreement that depth magnitude percepts increase with increasing disparity within Panum's fusional area, it is not clear whether performance follows geometric predictions and if not, what factors are responsible for the discrepancy. While previous research has shown that methods of manual depth estimation give comparable results to cue-comparison techniques when measuring perceived depth from motion parallax (Leonard, Nawrot, & Stroyan, 2013), to our knowledge there has been no comparison of intra- and cross-modal estimation methods in a single study, nor has there been a concerted effort to characterize the effect of experience on the pattern of depth estimates. Thus, the aim of this study is to consolidate and extend the existing knowledge concerning the perception of depth magnitude from binocular disparity, using observers with different degrees of expertise, and both intra- and cross-modal assessment methods.

2. Experiment 1

As discussed above, investigators have used cross-modal or intra-modal methods to estimate perceived depth to avoid the drawbacks associated with verbal reports and matching tasks. Generally, cross-modal techniques require that observers use the magnitude of sensation in one sensory modality to assess sensation in another modality. For example, Foley et al.'s (1975) manual pointing task is cross-modal because it requires that observers estimate depth magnitude (perceived visually) using a haptic response (e.g. pointing with an unseen finger). Such cross-modal techniques require a sensorimotor transformation from the visually perceived depth to a haptic response. In addition to the potential impact of memory in sequential estimates described above, this task requires the synchronization of hand-eye coordinates and potential reconstruction of the spatial interval (Anderson, Snyder, Li, & Stricanne, 1993; McGuire & Sabes, 2009). Digit span estimation tasks are also cross-modal in that observers are asked to use the distance between their thumb and index finger to estimate a displacement in depth. In both of these estimation methods, noise in the binocular disparity signal as well as the proprioceptive/motor system may influence the accuracy of depth estimates (Volcic, Fantoni, Caudek, Assad, & Domini, 2013). While it is impossible to eliminate all bias in cross-modal tasks, of these two, the digit span task is preferable because it avoids the under-reaching biases discussed by Foley and colleagues.

Unlike the cross-modal tasks described above, intra-modal depth estimation techniques rely on a transformation from disparity to depth that occurs within single sensory modality (Stevens, 1975). For example, Foley (1970) asked observers to adjust the position of a light point to represent half or twice the distance between a fixed reference and a target. While this task required that observers make a motor response (i.e. button press) the target-response transformation was within a single, visual modality. Normally, intra-modal estimation techniques also require a spatial transformation. For instance, in stereoscopic depth estimation tasks the target and reference stimuli are displaced along the z-axis, which is orthogonal to the fronto-parallel plane (x-axis). The comparison stimulus is then adjusted within this fronto-parallel plane. It has been noted that mental rotation operations needed to make this type of judgement may be subject to individual differences in spatial ability (Khooshabeh & Hegarty, 2010).

The tasks described above are subject to yet another potential source of bias or variability that stems from individual differences in experience (Foley & Richards, 1974; McKee & Taylor, 2010). Like many visuospatial abilities, studies of stereoacuity have shown that performance is highly dependent on the observers' experience with the stimuli and task; with focussed and prolonged training, performance can improve markedly (Fendick & Westheimer, 1983). However, the amount of improvement can vary widely across observers resulting in substantial interobserver variability (McKee & Taylor, 2010; Schmitt, Kromeier, Bach, & Kommerell, 2002). In our first experiment, we tested two groups of observers; one group had extensive experience with stereoscopic stimuli displayed on computer screens in a modified Wheatstone arrangement, while the other had no prior experience with either this type of stimuli or psychophysical tasks in general.

2.1. Methods

2.1.1. Observers

Eight experienced stereoscopic observers (including one author) were recruited. These observers had excellent stereoacuity