Optimal Combination of the Binocular Cues to 3D Motion

Brian Allen, Andrew M. Haun, Taylor Hanley, C. Shawn Green, and Bas Rokers

Department of Psychology, University of Wisconsin-Madison, Madison, Wisconsin, United States

PURPOSE. Perception necessarily entails combining separate sensory estimates into a single coherent whole. The perception of three-dimensional (3D) motion, for instance, can rely on two binocular cues: one related to the change in binocular disparity over time (CD) and the other related to interocular velocity differences (IOVD). Although previous work has shown that neither cue is strictly necessary for the perception of 3D motion, observers are able to judge 3D motion in displays in which one or the other cue has been eliminated, it is unclear whether or how the two cues are combined in situations in which both are present.

METHODS. We tested the visual performance of a sample of 81 individuals ($M_{age} = 20.34$, 49 females) in four main conditions that measured, respectively, static stereoacuity, CD, IOVD, and combined CD+IOVD sensitivity.

RESULTS. We show that the sensitivity to the two binocular cues to 3D motion varies substantially across observers (CD: $M_{d'} = 1.01$, $SD_{d'} = 1.1$; IOVD: $M_{d'} = 1.16$, $SD_{d'} = 1.03$). Furthermore, sensitivity to the two cues was independent across observers ($r[48] = 0.12$, $P = 0.42$). Importantly, however, observed CD+IOVD performance was well-predicted based on the assumption that each observer combines the two cues in a statistically optimal fashion ($r[79] = 0.75$, $P < 0.001$).

CONCLUSIONS. Our findings provide an explanation for the previously puzzling variability found in 3D perception across observers and laboratories, with some results suggesting that motion-in-depth percepts are largely determined by changes in binocular disparity, whereas others indicate that interocular velocity differences are key. Our results underline the existence of two complementary binocular mechanisms underlying 3D motion perception, with observers relying on these two mechanisms to different extents depending on their individual sensitivity.

Keywords: 3D motion, interocular velocity differences, changing disparity, cue combination

Our sensory systems provide us with a host of independent measurements about objects in the world. One of the key challenges of the perceptual system is thus to rationally combine these disparate estimates into a reasonable whole. This is true whether the estimates arise via different sensory modalities (e.g., combining auditory and visual estimates of an object's location[1]) or come from a single sensory modality (e.g., combining monocular and binocular visual estimates of the slant of a surface[2]). In the case of three-dimensional (3D) motion, in addition to a number of monocular cues, there are two binocular cues that can contribute to perception: changing binocular disparities and interocular velocity differences.[3] Under natural viewing conditions, changing disparity (CD) and interocular velocity differences (IOVD) tend to co-occur, with the primary functional difference arising due to a difference in the order of operations. In CD, binocular disparity for a feature is computed first, followed by computation of the change in disparity over time; in IOVD, change in monocular feature position is computed first, and the difference in velocity is computed subsequently[3,4] (cf. fig. 1 of Nefs et al.[5]; also fig. 1 of Peng and Shi[6]).

The relative importance of these binocular cues to 3D motion perception has been debated in recent years, with some researchers claiming that 3D motion perception largely depends on disparity-based cues,[4–7] whereas others have argued that there is a considerable role for velocity-based cues.[8–13]

Neurophysiological studies have not yet been able to adjudicate between the possibilities. Indeed, it is not obvious from the neurophysiology what course the visual system takes, as both binocular disparity and monocular direction are processed at multiple cortical sites.[14–21] For example, both types of motion-in-depth information seem to be processed in cortical area hMT+,[20] with IOVD being the main driver of motion-in-depth selectivity.[21] At least one study reports CD signals may be processed in a cortical area directly anterior to hMT+ in the later occipital complex.[17]

Our goal here was thus to identify the contribution of both binocular cues to 3D motion perception and determine if and how these cues are combined. We therefore tested the visual performance of a large sample of individuals in four main conditions that measured, respectively, static stereoacuity, CD, IOVD, and combined CD+IOVD performance. The relationship between CD and IOVD sensitivity both within and between observers demonstrates the degree to which these cues are processed independently. Furthermore, the relationship among CD, IOVD, and combined CD+IOVD sensitivity gives insight into how these cue sensitivities are used when a stimulus contains both cues (as is typically the case).
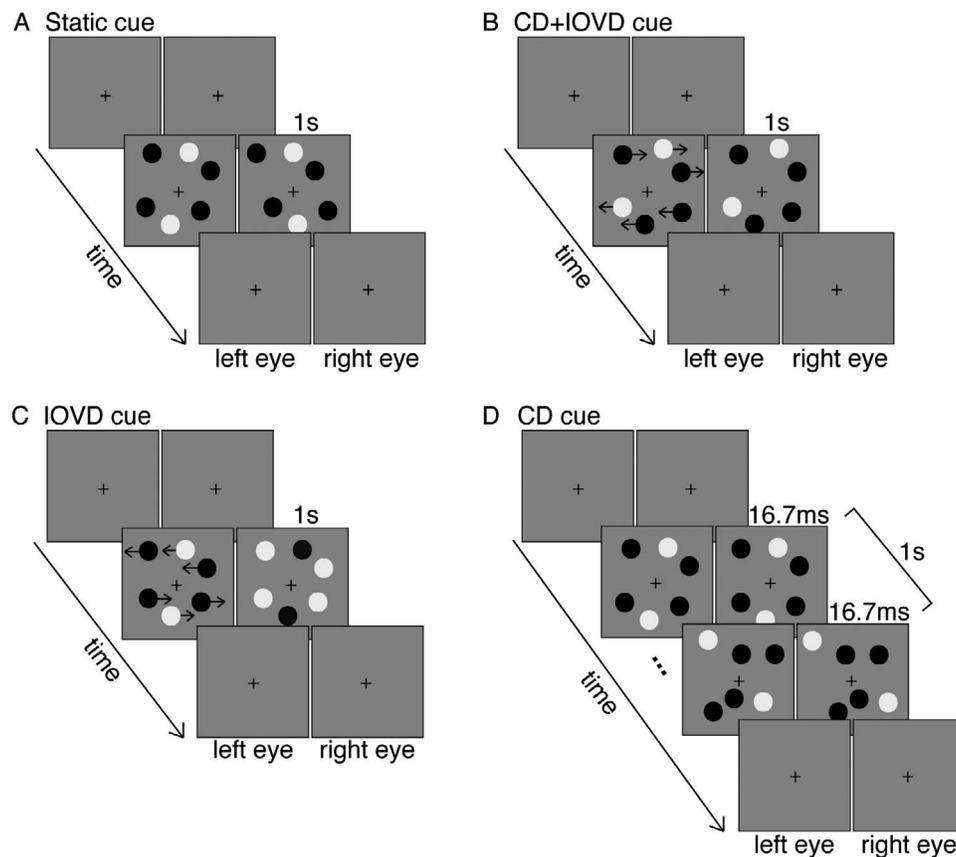
Investigative Ophthalmology & Visual Science

FIGURE 1. Schematic of the 3D stimuli. (**A**) The static cue stimulus tested static disparity perception. (**B**) The CD+IOVD cue stimulus tested both interocular velocity differences and changing disparity perception. (**C**) The IOVD cue stimulus tested ability to use velocity differences in the two eyes to infer the motion through depth of random dots. (**D**) The CD cue stimulus tested ability to use changing disparity to infer the motion in depth of random dot displays.

METHODS

Participants

A total of 81 participants (49 females), aged 18 to 51 (M_{age} = 20.34, SD_{age} = 5.92), completed the study. Participants were recruited from the University of Wisconsin-Madison campus and received extra credit for introductory psychology courses as compensation. All participants wore their normal prescription if any. Informed consent was obtained in accordance with the Declaration of Helsinki and the requirements of the institutional review board committee of the University of Wisconsin-Madison.

Apparatus

All visual tasks were performed on a Quad Core Intel Mac Pro (Apple, Inc., Cupertino, CA, USA) with an NVIDIA Quadro 4000 GPU (NVIDIA, Santa Clara, CA, USA), running Matlab (The MathWorks, Inc., Natick, MA, USA) and the Psychophysics Toolbox.[22,23] Visual stimuli were presented on a 54.6×33.8-cm LCD display (Planar SA2311W; Planar, Beaverton, OR, USA) 120Hz, 1920×1080 pixels) at a viewing distance of 85 cm for 3D stimuli and 59 cm for the other visual tasks (see Additional Measures, below). Participants wore active stereo shutter glasses (NVIDIA 3D 2, 60 Hz/eye), through which they viewed the LCD display. When viewed through the shutter glasses, the luminance of a white stimulus was 13.8 cd/m², midgray was 5.8 cd/m², and black was 0.03 cd/m². Head motion was minimized with the use of a chin rest.

Stereovision Tasks

Participants performed four tasks designed to measure their static and dynamic stereovision. Each block took approximately 5 minutes to complete and consisted of 100 trials. See Figure 1 for schematics of the different stimuli.

Static. For the static 3D stimulus, participants fixated the center of a Nonius cross at the center of the screen while two arrays of randomly positioned black and white dots (128 dots total) were presented simultaneously above and below fixation for 1 second on a midgray background. Each array extended from 0.5 to 6 degrees of visual angle above and below fixation and was 13 degrees wide (Fig. 1A). On each trial, one of the arrays was randomly selected to appear behind the plane of fixation (farther away), while the other array was presented in front of it (nearer). Total disparity was ±0.125 degrees. A 1/f (pink) noise pattern was presented in the spatial surround. Participants used the up or down keys to indicate which dot array (top or bottom) appeared nearer. The total disparity of 0.125 degrees was set according to pilot testing, so that the average performance across participants was approximately 75% correct.

Dynamic. We assessed sensitivity to 3D motion by using three versions of a dynamic 3D stimulus in which specific cues to 3D motion (changes in disparity and interocular velocity) could be isolated. In all stimuli, configuration of the display was similar to that described above for the static condition (extent, distribution, and contrast of dots), with the exception that the dots in the two arrays moved, indicating opposite

directions of motion-in-depth (toward and away from the observer). On the first frame of each trial, one of the arrays was randomly selected to appear behind the plane of fixation while the other array was presented in front of it (at 0.125 degrees of crossed/uncrossed binocular disparity). The arrays moved in opposite directions in depth at a speed of 0.25 degrees per second for 1 second, so that one array started 0.125 degrees in front of the plane of fixation and receded to 0.125 degrees behind the plane of fixation (and vice versa for the opposite array) on each trial. The array of dots that was presented behind fixation always approached and the array presented in front of fixation receded. Participants reported which dot array appeared to move toward them.

Changing Disparity Cue Stimulus. To isolate the CD cue to motion-in-depth (i.e., to remove interocular velocity differences), dots were continuously repositioned while gradually changing their binocular disparity over time. Perceptually the stimulus appears as a plane of detuned television snow moving through depth. With each screen refresh, the dot disparity was increased/decreased (depending on the direction of motion-in-depth of the given array) so that the disparity of the dots changed at a rate of 0.25 deg/s. In such a stimulus, individual dots do not provide a motion signal, but as a whole the changing binocular disparity of the dots defines a plane that moves through depth. Accuracy in this task thus provides a measure of sensitivity to changes in stimulus disparity over time (see Supplementary Movies for illustration of the 3D motion stimuli).

Interocular Velocity Difference Cue Stimulus. To isolate the interocular velocity difference cue (i.e., to attenuate information about changes in disparity), dots were given opposite contrast in each eye (i.e., black in one eye, white in the other). Although this does not entirely remove information about changes in disparity (CD is a necessary correlate of IOVD, but not vice versa), anticorrelation of stereo image pairs has been shown to significantly reduce the ability to use disparity information to perceive depth.[7,24,25] Accuracy in this task provides a measure of sensitivity to the differential direction of movement of a stimulus in each eye.

Combined Cues Stimulus. The combined cues (CD+IOVD) task block contained static disparity, changing disparity and interocular velocity cues, consistent with what would be present in natural viewing conditions. This task provides a general measure of sensitivity to the direction of motion-in-depth of a stimulus.

Additional Measures

We also conducted a visual test battery that included tests of visual acuity and speed of processing. These measures were taken to control for individual differences in visual ability that are not directly related to stereo processing.

Temporal Order Judgment

Because computer-presented motion stimuli are technically apparent motion stimuli, it is possible that individual differences in ability to detect changes in spatial-temporal sequencing could be a confounding factor in participants' ability to use monocular apparent motion information during a binocular task. One way individual variability in temporal order judgment is measured is with stimulus onset asynchrony (SOA) tasks.[26,27] Accordingly, an SOA task was used to measure temporal order discrimination. During each trial, participants fixated a central point (a 1° white rectangle against a midgray background) while two circles (diameter of 1°) appeared at slightly different times 5° above and below the fixation point. We varied the onset differences between 8 and 342 ms. Participants reported

which circle appeared first by using the up and down arrow keys on a standard keyboard. After each trial, participants received feedback about whether or not they responded correctly. A short (approximately 30-second) practice was completed before the task. The main task took approximately 5 minutes to complete and consisted of 12 trials per onset speed (for a total of 60 trials). For data analysis purposes, performance on the task was reduced to the linear slope of the fit of the participant's performance across the five SOA levels.

Speed of Processing

Because the 3D stimuli are presented in relatively short windows (1 second), another potentially confounding factor in participants' performance in the 3D discrimination tasks is individual differences in ability to rapidly process visual stimuli. Studies have shown fair variability in speed of processing measures (often response time) across participants and stimuli.[28,29] Accordingly, a simple discrimination task was used as an additional measure of the participants' speed of visual processing. During each trial, as participants fixated the center of the screen, either a white square or circle would appear (subtending 2° of visual angle). Participants were instructed to respond as fast as they could whether a circle or a square appeared by using the left and right arrow keys (left for square, right for circle) on a standard keyboard. After each trial, participants were told whether or not they responded correctly as well as their reaction time. Mean response time was used as the measure of simple discrimination abilities. A short (approximately 30-second) practice was completed before the experimental task. The task took roughly 4 minutes to complete and consisted of 120 trials.

Acuity

A Tumbling E task was used to measure participants' visual acuity at 5° and 15° of eccentricity (measured in separate blocks), which provides a measure of peripheral acuity. During the task, an "E" appeared either to the left or right of fixation, at which point the participant responded at which direction the E was facing by using the arrow keys (four cardinal directions). After each trial, participants received audio feedback as to whether or not they answered correctly. The stimulus size was controlled via a 3:1 staircase (i.e., after three correct responses the stimulus was reduced in size, after one incorrect response the stimulus was increased in size). The stimulus was changed by 50% during the first 20 trials, by 30% for the next 20 trials, and by 20% for the final 40 trials (80 trials in total). The task at each eccentricity (5° and 15°) took approximately 4 minutes. A short practice (approximately 30 seconds) was completed before the experimental Tumbling E task (at 5 degrees).

Relation to Stereo Measures

The measures of visual acuity, speed of processing, and SOA were not significantly correlated with any of the stereo measures (all $P > 0.30$).

Before the stereo experiment, participants completed 20 practice trials of the CD+IOVD cue condition, with audio feedback on whether or not they answered correctly. The practice trials used the combined CD+IOVD cue condition so that participants had equal prior experience with all cues. Participants always completed the CD+IOVD cue stimulus block next; the order in which participants completed the other three conditions (static, CD, and IOVD) was randomized among participants.
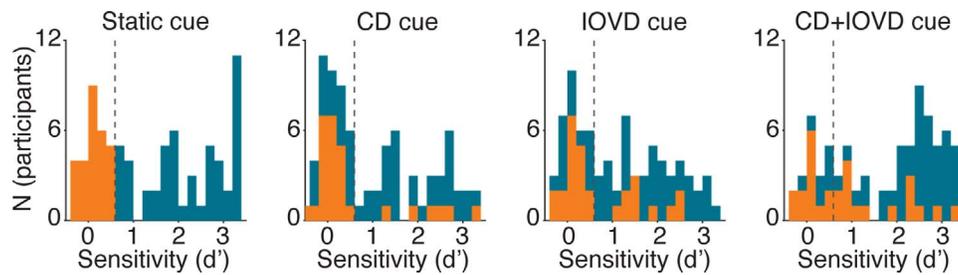
FIGURE 2. Distribution of performance across the four 3D tasks. *Dashed line* indicates level beyond which performance significantly differed from chance (at the 99% confidence level). *Orange bars* indicate observers categorized as stereo-anomalous based on performance in the static disparity task.

Quantifying Sensitivity

We estimated observer sensitivity by computing d' as the z-score of hit rate (upper array moved toward, observer reported "up") minus the z-score of false-alarm rate (lower array moved toward, observer reported "up"), divided by $\sqrt{2}$. We adjusted hit rates of 100% down to the next highest possible score (99%) in accordance with a 1/2N adjustment (see Ref. 30, under "general comments"). Likewise, we adjusted 0% false-alarm rates up to the next lowest possible score (1%).

RESULTS

"Stereo-Anomalous" Participants

Stereo-blindness is defined as an inability to see depth from disparity (i.e., a lack of stereopsis).[31,32] It is unknown whether stereo-blind individuals fall at the tails of the normal stereo-vision distribution or if they constitute a separate group, so we considered the data set from both perspectives. Accordingly, we set a "stereo-anomaly" threshold: we classified as "stereo-anomalous" any participant whose sensitivity on a stereovision condition was less than $d' = 0.59$, the maximum value at which the 99% confidence interval includes $d' = 0$, and was thus not significantly different from chance at the 99% confidence level.[33] We use the term "anomalous" instead of "blind" to describe these participants because we are cautious of claiming they cannot see depth from disparity per se (see Discussion).

Overall, of 81 participants, 30 (37%) showed below threshold sensitivity to the static cue stimulus, 42 (52%) showed below-threshold sensitivity to the CD cue stimulus, 32 (39%) showed below-threshold sensitivity to the IOVD cue stimulus, and 20 (25%) showed below-threshold sensitivity to the combined CD+IOVD cue stimulus (see Fig. 2 for histograms). Thirteen participants (16%) showed below-threshold sensitivity to *each* of the four stereovision stimuli (conversely, 28 [35%] participants showed above or at threshold sensitivity to all four stereo stimuli); 24 participants (30%) showed below-threshold sensitivity on average (averaging across the four stimulus conditions).

In pilot experiments, we observed considerable variability in performance across observers on these tasks. Our analysis hinges on the ability to compare an individual's performance across the stimulus conditions, rather than assessing absolute performance across the population. Consequently, to avoid large ceiling or floor effects, we set stimulus parameters across these tasks such that average performance was approximately 75% correct. Although this results in relatively high proportions of estimated stereo-anomaly for each individual stimulus, especially when compared with typical reports of stereo-blindness in the healthy population (1%–14%),[31,34–38] we believe there are a number of factors in our experiment (such

as stimulus density, and the 1-second trial length) that make this task considerably more difficult for naïve participants.

In addition to analyzing the sensitivity of all participants to these stimuli, "stereo-anomalous" and "stereo-normal" participants were considered in separate analyses, because their differential performance might bias the relationship between related tasks. Separating participants based on their sensitivity relative to threshold was also motivated by a previous study in which significant relationships between CD and IOVD sensitivity was observed when breaking up participants into two groups based on above- or below-average performance.[5] Unless otherwise stated, "stereo-normal" refers to participants whose static 3D sensitivity was measured as greater than or equal to $d' = 0.59$. Likewise, "stereo-anomalous" participants are those whose d' during static 3D task was less than the 0.59 threshold.

Does Static Depth Perception Limit/Predict 3D Motion Perception?

We first asked to what extent 3D motion perception could be predicted by (static) depth perception. We might expect that individuals who are stereo-anomalous for static stimuli should also be blind to motion-in-depth, reasoning that 3D motion is computed on the basis of previous computations of binocular disparity.[5,17]

Across all participants, sensitivity to all three dynamic 3D stimuli significantly and positively correlated with sensitivity to the static 3D stimulus (all $r[79] > 0.48$, $P < 0.001$) (see Fig. 3). In "stereo-normal" participants, sensitivity to the three dynamic 3D stimuli was significantly and positively correlated with sensitivity to the static 3D stimulus (all $r[49] > 0.50$, $P < 0.001$); however, in "stereo-anomalous" participants, no significant correlations were observed between static 3D stimulus sensitivity and sensitivity to the three dynamic 3D stimuli (all $r[28] < 0.13$, $P > 0.5$). Although inspection of the pattern of responses in Figure 3 similarly suggests no clear relationship between static cue sensitivity and sensitivity to the three dynamic stimuli for "stereo-anomalous" participants, it should be noted that the range of static cue sensitivity (d') values is, by definition, small for these participants, which has the effect of reducing the power of any test of statistical correlation.

Relationship Between Sensitivities to 3D Motion Stimuli

Across all participants, all pairwise correlations of sensitivity (d') to the three motion-in-depth stimuli were significantly positively correlated (all $r[79] > 0.58$, all $P < 0.001$) (see Fig. 4). In "stereo-normal" participants, sensitivity to each of the two cue-isolating stimuli (CD and IOVD) was significantly
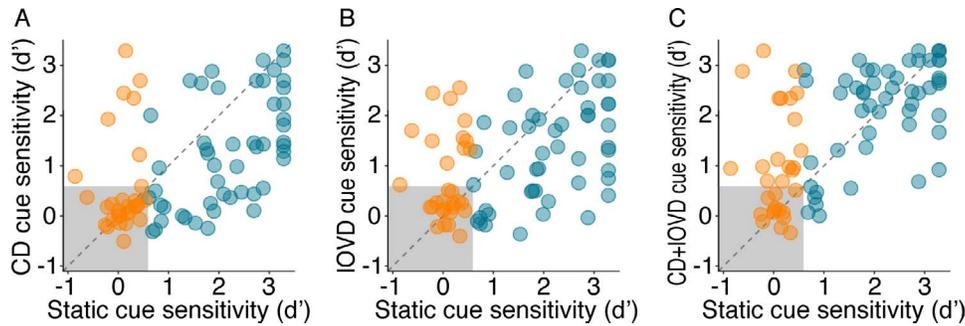
FIGURE 3. Correlations between sensitivity to the three dynamic stereovision stimuli and the static stereovision stimulus. (**A**, **B**) Relationship between sensitivity to the CD and IOVD cues and the static disparity cue, respectively. Most observers showed significantly greater sensitivity to the static disparity stimulus than to the CD and IOVD cue stimuli in isolation. (**C**) Relationship between sensitivity to the combined CD+IOVD cue and the static disparity cue. In contrast to the performance in (**A**, **B**), most observers were more sensitive to the direction of motion in depth when both 3D motion cues were present, than the static disparity cue. *Orange symbols* represent observers classified as stereo-anomalous based on static cue sensitivity. *Black dashed curve* represents the identity line. *Gray squares* indicate the range of sensitivity in each task not significantly different from chance.

positively correlated with sensitivity to the combined CD+IOVD stimulus (both $r[49] = 0.62$, $P < 0.001$).

Although overall we find a significant *positive* relationship between sensitivity to the CD and IOVD cues (Fig. 4A), the positive correlation between the two cues seems to be predominantly driven by the cluster of observers that did not perform significantly different from chance on either cue (Fig. 4A: data in the gray box). When removing participants who show below-threshold sensitivity to both the CD and IOVD stimuli (e.g., only considering data points that lie outside of the gray box in Fig. 4A), there is no longer a significant correlation between sensitivity to the CD and IOVD cues ($r[48] = 0.12$, $P = 0.423$). Additionally, we observe no significant correlation between CD and IOVD sensitivity when considering only participants who perform above threshold on both the CD and IOVD tasks ($r[32] = 0.3$, $P = 0.081$). These findings suggest that sensitivity to the two cues in above-threshold observers may be largely independent and that the correlation in performance is driven in large part by observers who perform poorly using both the CD and IOVD cues to 3D motion.

Finally, it is clear from Figure 4 that sensitivity to the combined CD+IOVD cue stimulus is significantly greater than sensitivity to either cue in isolation (most data points in Figs. 4B, 4C lie above the [the dashed gray diagonal] identity line). This strongly suggests that most observers combine the two

binocular cues in their judgment of 3D motion. In the next section, we will consider 3D motion perception as the optimal combination of the two cues in more detail.

Predicting Full-Cue Performance

We next asked whether and how sensitivity to "combined-cue" motion-in-depth perception depends on its IOVD and CD components. Optimal (or near optimal) combination of independent cues is often observed in other perceptual contexts.[39–43] If the two cues make independent contributions, the optimal combination of two signal-to-noise ratios (d' values) is their Euclidean sum,[44] so we computed these as "predicted full-cue sensitivities," and compared them to the observed full-cue performance (Fig. 5B).

$$d'^2_{IOVD+CD} = d'^2_{IOVD} + d'^2_{CD}$$

We found that the observed full-cue sensitivity was correlated with the predicted (optimal) full-cue sensitivity ($r[79] = 0.75$, $P < 0.001$, see Fig. 5B) (with "predicted" sensitivities capped to the maximum measurable value in our experiment: d' = 3.29). To estimate more precisely whether or not our observers tended to combine IOVD and CD cues optimally, we measured the previously proposed "integration index"[45]:
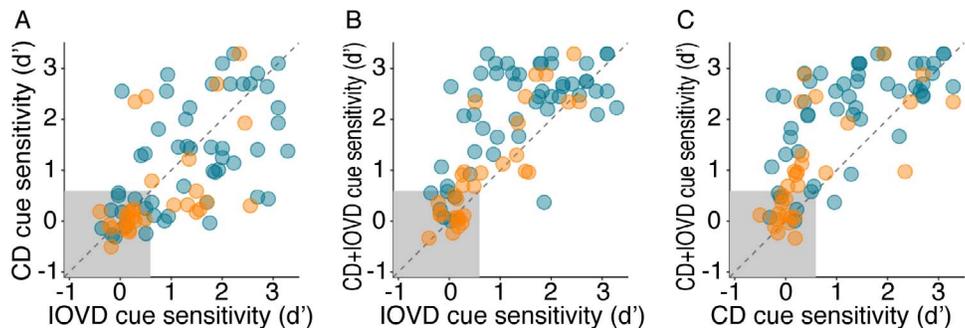

FIGURE 4. Relationship between sensitivity in isolated and combined-cue 3D motion conditions. *Blue symbols* indicate "stereo-normal" observers, and *orange symbols* indicate "stereo-anomalous" observers. (**A**) When considering all observers, CD and IOVD cue sensitivities are positively correlated. However, when considering only participants who show above-threshold sensitivity to at least one of the CD and IOVD cues (e.g., *data points* that lie outside of the *gray box*), there is no significant correlation between sensitivity to CD and IOVD cues, suggesting that the positive correlation in performance is driven in large part by observers who perform poorly using both the CD and IOVD cues to 3D motion, and that performance based on either cue in above-threshold observers is largely independent. (**B**, **C**) For the vast majority of subjects, stereo-anomalous or not, sensitivity to the combined CD+IOVD stimulus exceeds sensitivity to either cue in isolation (*data points* lie above the *positive diagonal*), indicating observers combine the two cues to 3D motion.
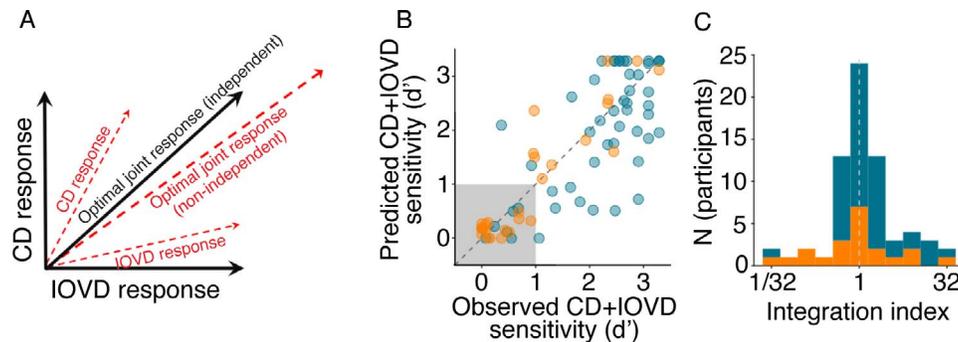
FIGURE 5. Assessment of optimal combination of independent cues to 3D motion. (**A**) Illustration of optimal cue combination. The optimal joint response under an assumption of cue independence is the vector sum (*diagonal black arrow*) of the individual cue contributions (*horizontal/ vertical arrows*). In case the cues are nonindependent (*nonorthogonal dashed red arrows*) the optimal joint response will be larger (superoptimal) and predict a larger joint response (*diagonal red arrow*). (**B**) Relationship between "optimal" predicted performance and observed full (CD+IOVD)-cue performance assuming independent cue contributions. *Orange symbols* are "stereo-anomalous" subjects whose d′ in the static disparity task was less than 0.59; *blue symbols* are "stereo-normal." (**C**) Distribution of integration index (distance from the *diagonal* in [A]) in 65 subjects, excluding those for whom sensitivity was less than or equal to zero in both of the single-cue tasks. The distribution is narrowly centered around 1, suggesting that observers optimally combine the independent CD and IOVD cues to 3D motion.

$$I = \frac{d'^2_{IOVD+CD}}{d'^2_{IOVD} + d'^2_{CD}}.$$

This index has a value of 1 if combination is optimal (Euclidean), less than 1 if combination is suboptimal (i.e., both cues cannot be fully used at once), and greater than 1 if combination is superoptimal (i.e., if the cues are not processed independently, as illustrated in Fig. 5A). We computed the integration index only for those subjects ($n = 65$) who had non-zero (>0) sensitivity to at least one single-cue condition (because the index is not computable otherwise) and whose combined-cue sensitivity was greater than zero. Figure 5C shows that the integration index was narrowly distributed around a value of 1.0, suggesting that the two binocular cues to 3D motion are indeed independent, and that observers seem to combine the two cues in a statistically optimal fashion.

DISCUSSION

We assessed the extent to which the perception of 3D motion-in-depth is driven by estimates of CD and/or estimates of IOVD. Specifically, we tested depth perception based on sensitivity to binocular disparity, and the ability to perceive 3D motion based on CD, IOVD, or the combination of both (CD+IOVD) cues. We did not find evidence that one cue or the other better predicted 3D motion perception across the population. Instead, we found considerable variability in the sensitivity to both cues across observers, such that some observers were more sensitive to disparity-based cues, whereas others were more sensitive to velocity-based cues. For example, a quarter of our participants (21 of 81) showed sensitivity significantly above chance for only one of the two dynamic cue-isolating 3D motion stimuli (CD and IOVD). We further found that when presented with visual stimuli that contained both cues to 3D motion, observers tended to be more sensitive to the direction of motion-in-depth than would be predicted based on their sensitivity to either cue alone. A model that assumed that the cues are processed independently and are optimally combined according to their reliability provided a good fit to the data. These results should help clarify the sometimes inconsistent findings across previous studies,[4–13] suggesting that previous results may have been due to natural variability in the population, and the small sample sizes typical of intensive psychophysics.

Relationship Between Position-in-Depth and Motion-in-Depth Mechanisms

Although in the aggregate, participants tended to be more sensitive to the IOVD stimulus, we observed several participants who were sensitive only to the CD stimulus and who failed to show above-threshold sensitivity to IOVD. Perhaps somewhat surprisingly, we also found a number of participants who were not sensitive to static disparity task, but were to one or more dynamic 3D stimuli, including the dynamic disparity (CD) task. A potential explanation for this seemingly contradictory pattern of sensitivities is that these participants might not have been sensitive to the exact fixed binocular disparity of the static stimulus (7.5 min arc) but could detect a change through the range of disparities (centered on the static disparity value) in the dynamic task.

It should be noted that the 1-second stimulus presentation durations may not have been optimal for the static disparity task, because stereo mechanisms are generally considered to be slow, and participants therefore may have performed better at longer durations. At the same time, however, participants clearly show an ability to process binocular information over the 1-second time course in the motion-in-depth tasks, and consistent with an upper limit on integration, recent work has suggested that motion-in-depth performance ceases to improve with durations over approximately 1 second.[46]

Relationship to Previous Studies on Stereo-Anomaly

Most psychophysical investigations into stereo-blindness and stereo-anomaly report stereo-blindness in between 1% and 14% of participants.[31,34–38] A recent, carefully controlled large-cohort investigation[38] reports stereo-blindness in 2.2% of their participants. These accounts are markedly lower than the 37% of participants we classify as "stereo-anomalous" for static disparity in the current study. This discrepancy can likely be explained by the wide variability in criteria, stimuli, subject training, and task properties across these studies. For instance, to ensure that the different stimulus conditions were equated, we used the same disparity ranges and dot densities across all conditions, and limited stimulus presentation time to 1 second. Conversely, typical clinical assessments of stereo acuity provide much longer or even unlimited viewing time. Large improvements in stereo test performance after "encouraging"

participants to "tune in" to the stimulus have been reported.[38] On the other hand, psychophysical assessments of stereo-acuity thresholds[47–49] typically produce thresholds less than 1 minute of arc, and one might thus have expected performance to be at ceiling for all our non–stereo-blind observers. However, we would like to emphasize that such psychophysical studies typically rely on a small number of highly experienced observers, and it is known, although unfortunately not frequently reported, that performance for truly naïve observers can initially fall well short of such performance levels. Indeed, in initial piloting, using disparity values and stimulus presentation times informed by expert observer performance, we found near-floor performance for the vast majority of naïve observers.

Although we did expose participants to 20 training trials (in which they received feedback) with the combined CD+IOVD cue stimulus before the experiment, we cannot be sure this feedback helped "tune in" their stereovision. We do not observe a significant increase in performance when comparing the first and last 50 trials of the CD+IOVD blocks across observers, suggesting significant perceptual learning did not occur during this task. Accordingly, because performance on stereo (3D) tasks strongly depends on threshold criteria as well as task and stimulus properties, we opted to use the term "stereo-anomalous" rather than stereo-blind in our current study and focused on the variation of performance across the population.

We did not explicitly assess monocular motion perception. No observer reported an inability to see motion and there are no reports in the literature of observers being unable to see the direction of monocular motion in fully coherent displays as used here. This means that any inability to perceive the direction of motion in depth is specific to impairment in combination of the monocular motion signals. The neural origin of this impairment remains poorly understood.

Implications for Motion-in-Depth Mechanisms

In our sample, sensitivity to CD and IOVD cues was uncorrelated when considering participants whose thresholds were significantly above chance on both tasks, leading us to conclude that these cues are processed via largely independent mechanisms. This runs contrary to a previous report[5] of an inverse relationship in sensitivity to the CD and IOVD cues across observers, suggesting that this reflected a "fallback" for participants with poor stereo perception (i.e., such observers might resort to using IOVD cues in place of CD).

It is possible that the incongruence of our findings with those previously reported[5] arises from differences in the stimuli used. The IOVD stimulus in the current study used "anticorrelated" dots: dots whose physical positions are matched between the two eyes as they move but are opposite in contrast. Such anticorrelated stereo displays are known to *not* evoke stereoscopic percepts,[24,25] although they may still evoke disparity-selective responses in some visual neurons.[50] The IOVD stimulus used in previous work[5] used "uncorrelated" dots: dots whose physical positions are not matched between the eyes and move in opposite directions in each eye. The IOVD stimulus in our study thus eliminates perceptible stereo information, but allows some disparity information to be processed. This information might conceivably have contributed to a (weak) positive relationship between CD and IOVD in the current study negating the negative correlation. However, previous neuroimaging research has demonstrated differential activity in hMT+ when viewing similar dynamic anticorrelated 3D stimuli versus IOVD-containing 3D stimuli,[19] suggesting these anticorrelated CD displays are processed in a measurably different way from similar IOVD stimuli in hMT+.

It has also been suggested previously that in the presence of two detectable cues, observers default to one "preferred" cue.[5] By contrast, we found that performance in the combined CD+IOVD cue condition almost always exceeded performance based on either cue in isolation, and that our observers combined CD and IOVD information in a relatively optimal fashion, although in some cases we observed a trend toward superoptimal combination in the stereo-normal subjects ("integration index" greater than 1; Fig. 5C).

Implications for Clinical Testing

Several of our participants who were deemed "stereo-anomalous" based on their sensitivity to the static 3D stimulus nonetheless showed above-threshold sensitivity 3D motion, including stimuli that solely contained changes in disparity. This highlights the need for more careful evaluation of stereovision abilities when classifying participants based on psychophysical performance. Furthermore, future psychophysical investigations into abnormal binocular function warrant careful testing of stereovision abilities, including tests of sensitivity to 3D motion.

CONCLUSIONS

In summary, the results of the current study help reconcile the inconsistent reports on 3D motion perception reported in the literature. Understanding how visual cues are integrated is critical for uncovering the mechanistic basis of neural-based visual disorders affecting binocular integration, such as amblyopia, and its consequences for impaired function in motion-in-depth perception.

References

1. Shams L, Kamitani Y, Shimojo S. Visual illusion induced by sound. *Brain Res Cogn Brain Res*. 2002;14:147–152.

2. Knill DC, Saunders JA. Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Res*. 2003;43:2539–2558.

3. Regan D. Binocular correlates of the direction of motion in depth. *Vision Res*. 1993;33:2359–2360.

4. Nefs H, Harris J. What visual information is used for stereoscopic depth displacement discrimination? *Perception*. 2010;39:727–744.

5. Nefs H, O'Hare L, Harris J. Two independent mechanisms for motion-in-depth perception: evidence from individual differences. *Front Psychol*. 2010;1:155.

6. Peng Q, Shi B. Neural population models for perception of motion in depth. *Vision Res*. 2014;101:11–31.

7. Cumming B, Parker A. Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature*. 1997;389:280–283.

8. Brooks K. Monocular motion adaptation affects the perceived trajectory of stereomotion. *J Exp Psychol Hum Percept Perform*. 2002;28:1470–1482.

9. Fernandez J, Farrell B. Motion in depth from interocular velocity differences revealed by differential motion aftereffect. *Vision Res*. 2006;46:1307–1317.

Investigative Ophthalmology & Visual Science

10. Shioiri S, Nakajima T, Kakehi D, Yaguchi H. Differences in temporal frequency tuning between the two binocular mechanisms for seeing motion in depth. *J Opt Soc Am A Opt Image Sci Vis*. 2008;25:1574–1585.

11. Shioiri S, Kakehi D, Tashiro T, Yaguchi H. Integration of monocular motion signals and the analysis of interocular velocity differences for the perception of motion-in-depth. *J Vis*. 2009;9(13):10.

12. Czuba TB, Rokers B, Guillet K, Huk AC, Cormack LK. Three-dimensional motion aftereffects reveal distinct direction-selective mechanisms for binocular processing of motion through depth. *J Vis*. 2011;11(10):18.

13. Sakano Y, Allison R. Aftereffect of motion-in-depth based on binocular cues: effects of adaptation duration, interocular correlation, and temporal correlation. *J Vis*. 2014;14(8):21.

14. Zeki SM. Cells responding to changing image size and disparity in the cortex of the rhesus monkey. *J Physiol*. 1974;242:827–841.

15. Maunsell J, Van Essen D. Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *J Neurophysiol*. 1983;49:1148–1167.

16. Albright T. Direction and orientation selectivity of neurons in visual area MT of the macaque. *J Neurophysiol*. 1984;52:1106–1130.

17. Likova L, Tyler C. Stereomotion processing in the human occipital cortex. *NeuroImage*. 2007;38:293–305.

18. Wallisch P, Movshon A. Structure and function come unglued in the visual cortex. *Neuron*. 2008;60:195–197.

19. Rokers B, Cormack LK, Huk AC. Disparity-and velocity-based signals for three-dimensional motion perception in human MT+. *Nat Neurosci*. 2009;12:1050–1055.

20. Czuba T, Huk A, Cormack L, Kohn A. Area MT encodes three-dimensional motion. *J Neurosci*. 2014;34:15522–15533.

21. Sanada T, DeAngelis G. Neural representation of motion-in-depth in area MT. *J Neurosci*. 2014;34:15508–15521.

22. Brainard DH. The psychophysics toolbox. *Spat Vis*. 1997;10:443–446.

23. Pelli DG. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis*. 1997;10:437–442.

24. Cogan A, Kontsevich L, Lomakin A, Hlpern D, Blake R. Binocular disparity processing with opposite-contrast stimuli. *Perception*. 1995;24:33–47.

25. Cumming B, Shapiro S, Parker A. Disparity detection in anticorrelated stereograms. *Perception*. 1998;27:1367–1377.

26. Kahneman D, Wolman RE. Stroboscope motion: effects of duration and interval. *Perception and Psychophysics*. 1970;8:161–164.

27. McKee SP. A local mechanism for differential velocity detection. *Vision Res*. 1981;21:491–500.

28. Potter MC, Faulconer BA. Time to understand pictures and words. *Nature*. 1975;253:437–438.

29. Thorpe S, Fize D, Marlot C. Speed of processing in the human visual system. *Nature*. 1996;381:520–522.

30. Macmillan N, Kaplan H. Detection theory analysis of group data: estimating sensitivity from average hit and false-alarm rates. *Psychol Bull*. 1985;98:185–199.

31. Richards W. Stereopsis and steroblindness. *Exp Brain Res*. 1970;10:380–388.

32. Lema SA, Blake R. Binocular summation in normal and stereoblind individuals. *Vision Res*. 1977;17:691–695.

33. Gourevitch V, Galanter E. A significance test for one parameter isosensitivity functions. *Psychometrika*. 1967;32:25–33.

34. Coutant B, Westheimer G. Population distribution of stereoscopic ability. *Ophthalamic Physiol Opt*. 1993;13:3–7.

35. Zaroff C, Knutelska M, Frumkes T. Variation in stereoacuity: normative description, fixation disparity, and the roles of aging and gender. *Invest Ophthalmol Vis Sci*. 2003;44:891–900.

36. Rahi J, Cumberland P, Peckham C. Visual impairment and vision-related quality of life in working-age adults: findings in the 1958 British birth cohort. *Ophthalmology*. 2009;116:270–274.

37. Bohr I, Read J. Stereoacuity with Frisby and revisited FD2 stereo tests. *PLoS One*. 2013;8:e82999.

38. Bosten J, Goodbourn P, Lawrance-Owen A, Bargary G, Hogg R, Mollon J. A population study of binocular function. *Vision Res*. 2015;110:34–50.

39. Landy M, Maloney L, Johnston E, Young M. Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Res*. 1995;35:389–412.

40. Ernst M, Banks M. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*. 2002;415:429–433.

41. Hillis J, Ernst M, Banks M, Landy M. Combining sensory information: mandatory fusion within, but not between, senses. *Science*. 2002;298:1627–1630.

42. Oruc I, Maloney L, Landy M. Weighted linear cue combination with possibly correlated error. *Vision Res*. 2003;43:2451–2468.

43. Faisal A, Wolpert D. Near optimal combination of sensory and motor uncertainty in time during a naturalistic perception-action task. *J Neurophysiol*. 2009;101:1901–1912.

44. Green DM, Swets JA. *Signal Detection Theory and Psychophysics*. New York: Kreiger; 1974.

45. Nandy A, Tjan B. Efficient integration across spatial frequencies for letter identification in foveal and peripheral vision. *J Vis*. 2008;8(13):1–20.

46. Katz L, Hennig J, Cormack L, Huk A. A distinct mechanism of temporal integration for motion through depth. *J Neurosci*. 2015;35:10212–10216.

47. Ogle K, Weil M. Stereoscopic vision and the duration of the stimulus. *AMA Arch Ophthalmol*. 1958;59:4–17.

48. Westheimer G, McKee S. What prior uniocular processing is necessary for stereopsis? *Invest Ophthalmol Vis Sci*. 1979;18:614–621.

49. Read J, Phillipson G, Serrano-Pedraza I, Milner A, Parker A. Stereoscopic vision in the absence of the lateral occipital cortex. *PLoS One*. 2010;5:e12608.

50. Neri P. A stereoscopic look at visual cortex. *J Neurophysiol*. 2005;93:1823–1826.