Critical Immaturities Limiting Infant Binocular Stereopsis

Angela M. Brown,[1,2] *Delwin T. Lindsey,*[2,3] *PremNandhini Satgunam,*[1,2] *and Jaime A. Miracle*[1]

Purpose. To determine what critical immaturity is responsible for the poor binocular stereopsis of human infants.

Methods. Infant and adult psychometric functions were measured for detection of stereoscopic depth in a random-texture display. A test stimulus defined by horizontal binocular disparity and a distracter stimulus defined by vertical disparity were used. Adults were tested by direct psychophysical methods at several contrast values, and infants by forced-choice preferential looking at 100% contrast.

Results. Infant stereoacuity matured from unmeasurable at age 12 weeks to 7.9 arc min at 20 weeks, which was still far from the nominal adult value of 5 to 10 arc seconds. In contrast, infant d-max (maximum disparity) was 86.8 minutes at 20 weeks, which was near the adult d-max of 110.6 minutes. The average maximum level of infant performance at 20 weeks was 77% correct, still far below adult performance. When the adult stereogram was low contrast, adult extrafoveal performance was similar to infant performance. Infant and adult stereo performance was predicted quantitatively, using infant and adult monocular performance in detecting the stereogram texture. Infant and adult stereopsis performance approached, but did not reach, the predicted values.

Conclusions. The infantlike performance of adults tested at low contrast and the similarity of infant maximum percentage of correct data relative to the predicted values suggested that the critical immaturity limiting infant stereopsis is the well-known insensitivity of the infant visual system to contrast. This conclusion supports the clinical use of stereopsis as a screening test for bilateral monocular function in infants. (*Invest Ophthalmol Vis Sci.* 2007;48:1424–1434) DOI:10.1167/iovs.06-0718

Binocular stereopsis is the ability of the observer to perceive relative distances in depth among objects in 3-D space, using the disparity between the right- and left-eye views of those objects as a cue. Stereopsis is a classic topic for research in infant visual development.[1–3] It has attracted this attention for two general reasons. First, basic scientists are interested in the development of stereopsis because (in adults) stereopsis is a property of visual function that depends critically on information processing that happens in the cerebral cortex (e.g., Ref. 4). The development of this cortical function is studied only if infant stereopsis is critically limited by the limited ability of the infant brain to compare the signals arriving from the two eyes and to extract a signal from the binocular disparity between them. Clinically, stereopsis is useful in diagnosing disorders specific to binocular function, such as strabismus and strabismic amblyopia. Second, clinicians are interested in testing stereopsis in infants as a screening tool: stereopsis is not possible unless both eyes generate useable visual signals.[5] Whereas testing one eye after the other necessarily confounds the relative performance of the two eyes with changes in the infant's state between the two tests, stereopsis allows the clinician to evaluate both eyes simultaneously. However, this second clinical use of stereopsis depends on the assumption that the infant's performance on a stereopsis task is limited by the ability of the infant to see the two hemistereograms simultaneously.

Thus, there are two possible critical immaturities that may limit an infant's stereopsis. On the one hand, visual processes specific to stereopsis may be critical.[6] In that case, the study of stereopsis may reveal immaturities or disorders of visual information processing in the cerebral cortex, but the use of stereopsis as a screening tool for bilateral monocular visual function should be re-examined. On the other hand, infant stereopsis may be critically limited by the well-known general insensitivity of infants to luminance contrast.[7] For example, the V1 cells of newborn macaques are sensitive to binocular disparity, but require higher contrast than adult V1 cells to function.[8] In that case, we may not learn very much about cerebral development from studying stereopsis, but the use of stereopsis in screening the monocular function of infant patients is well justified.

To decide between these alternative critical immaturities, we set out to study the early development of binocular stereopsis in infants. We collected data over a large range of binocular disparities in an effort to measure the minimum amount of disparity required for stereopsis (stereoacuity), the maximum amount of disparity that supports stereopsis (d-max), and the maximum level of stereo performance (in percentage correct) on a psychophysical task. We then used our results to evaluate the hypothesis that infant stereopsis is critically limited by the monocular visibility of the hemistereogram presented to each of the two eyes, against the alternative hypothesis that there is another "special" immaturity that prevents stereopsis, even for stimuli that are plainly visible to infants.

Experiment I: Infant Data

This was a behavioral study of binocular stereopsis in infants between 3 and 5 months of age. We chose a random stereo texture for our stereograms because we wanted to eliminate false matches that may have occurred if the hemistereograms were periodic (reviewed by Ref. 4) and to avoid the half-cycle limit that may have applied if the hemistereograms were modulated only in the horizontal dimension.[9–11] Stereograms were presented on a rear projection screen and viewed dichoptically by means of circular polarization to separate the two eyes'

From the [1]College of Optometry, the [2]Graduate Program in Vision Science, and the [3]Department of Psychology, The Ohio State University, Columbus, Ohio.

Supported by National Science Foundation Grant BCS-9983465 (AMB) and the Ohio Lions Eye Research Foundation, which also funded publication costs.

Submitted for publication June 27, 2006; revised August 24 and October 4, 2006; accepted January 19, 2007.

Disclosure: **A.M. Brown**, None; **D.T. Lindsey**, None; **P.-N. Satgunam**, None; **J.A. Miracle**, None

Corresponding author: Angela M. Brown, 338 W. 10th Avenue, Columbus, OH, 43210-1240; brown.112@osu.edu.

Investigative Ophthalmology & Visual Science, March 2007, Vol. 48, No. 3

stimuli. Each stereogram used horizontal binocular disparity to portray an 11.5° square, which protruded or receded in depth on one side of the screen. Each horizontal disparity target was paired with a distracter stimulus that was portrayed by an equal amount of vertical binocular disparity and was located on the opposite side of the screen. We chose a large square target and a low spatial frequency for the texture because those are the conditions that give the largest d-max in adults.[12] We chose a vertical disparity distracter to assure that any fixation preference for the test stimulus was based on its horizontal binocular disparity, rather than on some general preference for different images being presented to the two eyes. We also used a "flat" (binocularly identical) distracter stimulus on a few infants to assure that their poor overall performance was not due to the possibly "interesting" vertical disparity distracter stimulus. It is well-accepted that inaccurate vergence of the eyes does not limit stereopsis in infants,[13,14] not least because infant vergence is good over the age range where stereopsis is emerging.[15] However, this presents its own problems, and so we always mixed crossed and uncrossed disparity stimuli within blocks to minimize any tendency of the subject to fixate chronically nearer or farther than the base plane of the stimulus (a strategy suggested by Israel Abramov, Brooklyn College, CUNY, New York, NY). We assayed the fixation preference of our infant subjects using forced-choice preferential looking. The resultant psychometric functions were unimodal, and allowed estimation of the stereoacuity, the d-max, and the maximum level of detection performance.

Methods

Subjects. We recruited infant subjects by letter from the birth announcements in the local newspaper. Parents reported that their infants were born healthy within 3 weeks of their due dates and that no first-degree relatives had any history of strabismus, amblyopia or serious eye disorder. Testing began at any age between 84 and 153 days, and continued weekly for as many sessions as the parents wished or until the infant reached the age of 21 weeks. The parents of the infant subjects provided written, informed consent before testing began. All the research reported in this article adhered to the tenets of the Declaration of Helsinki and was approved in advance by the Biomedical Human Subjects committee of the Institutional Review Board of The Ohio State University. Parents were paid $10 for their infants' participation and were offered an optometric eye examination at the Pediatrics Service of The Ohio State University College of Optometry. Thirty-seven percent of our infants were examined, and no visual anomalies were noted except for clinically insignificant refractive errors.

Stimuli. The stereograms were created (Mathematica; Wolfram Research, Champaign, IL) and presented on computer in a slide show program (PowerPoint; Microsoft, Redmond WA). The hemistereogram intended for the right eye was projected via the red output channel of the computer, which drove the green primary of one video projector (DLA 2000; JVC, Wayne, NJ) and was projected via a right-handed circular polarizer onto a rear projection screen (ST-professional-W; ScreenTech, Myaree, Western Australia), which preserved polarization (crosstalk, 3.5%). The hemistereogram intended for the left eye was projected via the green output channel, which drove the green primary of a second, identical projector and was projected via a left-handed polarizer. The subject viewed the stereograms via a pair of pediatric spectacle frames, which were fitted with a right-handed circular polarizer over the right eye and a left-handed circular polarizer over the left eye. The use of circular polarizers ensured that the separation of the hemistereograms to the two eyes did not depend critically on the subject's head remaining perfectly upright. The space-average luminance of the display presented to each eye was 35 cd/m^2, as viewed through the polarizing glasses.

The stereogram textures were created from 440 × 440-pixel squares where each pixel was randomly assigned an initial value of either 1 or 0. The random pixels were filtered with a circular, symmetrical, tapered Bessell kernel with a peak spatial frequency of 0.95 cyc/deg. The filtered image was then thresholded at its median value, with half of its area being assigned to white and half to black, creating a stimulus with 100% Michelson contrast. The resultant texture is illustrated in Figure 1, and its spatial frequency spectrum is the unfiltered spectrum in Figure 2.

The large, square, textured stereograms appeared on the right and left sides of the screen (Fig. 1A). The test and distracter stimuli were smaller square regions (11.5° visual angle [VA] on each side), one within each of the larger patches of texture. Within the square region on the right or left of the screen, the texture presented to one eye was shifted to the right or to the left relative to the texture presented to the other eye, to define a test stimulus using crossed or uncrossed horizontal binocular disparity. Within the square region on the opposite side of the screen, the texture presented to one eye was shifted vertically by an equal amount, to define a distracter stimulus using vertical binocular disparity. There were texture discontinuities at the boundaries of the test and distracter stimuli so the binocular disparities could be small compared to the gauge of the texture. Thus, the boundaries and the horizontal binocular disparity portrayed the square test stimulus as either a card suspended in space in front of the base plane of the stimulus, or as a square aperture opening like a window onto a plane appearing behind the base plane of the stimulus. The boundaries were adjusted in position to portray the card or the window correctly. When we viewed the stimulus without the stereo glasses, we could not readily determine which side of the stimulus contained the horizontal binocular disparity; although, when the disparity was small, we could see it by scrutinizing the edges of the texture elements. In a control experiment, a few infants were tested with a "flat" distracter stimulus with the boundaries of the test stimulus in place, but with a vertical disparity of zero.

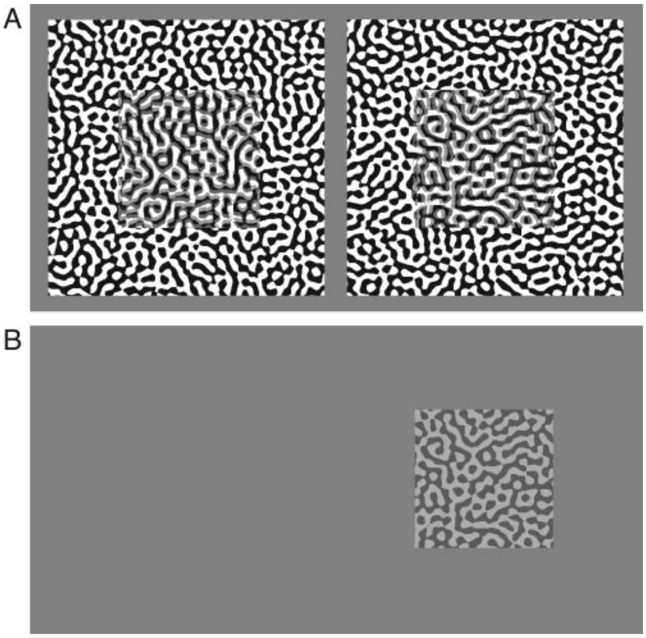

FIGURE 1. Typical 100% contrast stimuli, in gray tones. One eye's stimulus has been made lighter than the other, for clarity. In the actual experiments, binocular separation was maintained by circular polarization. **(A)** A stereo stimulus used in experiments I and II. **(B)** A monocular stimulus used in experiment III: in this example, the right eye sees a uniform field and the left eye sees a texture embedded in a uniform field. See Supplementary Figure S1, online at http://www.iovs.org/cgi/content/full/48/3/1424/DC1, for an anaglyphic representation of these stimuli.

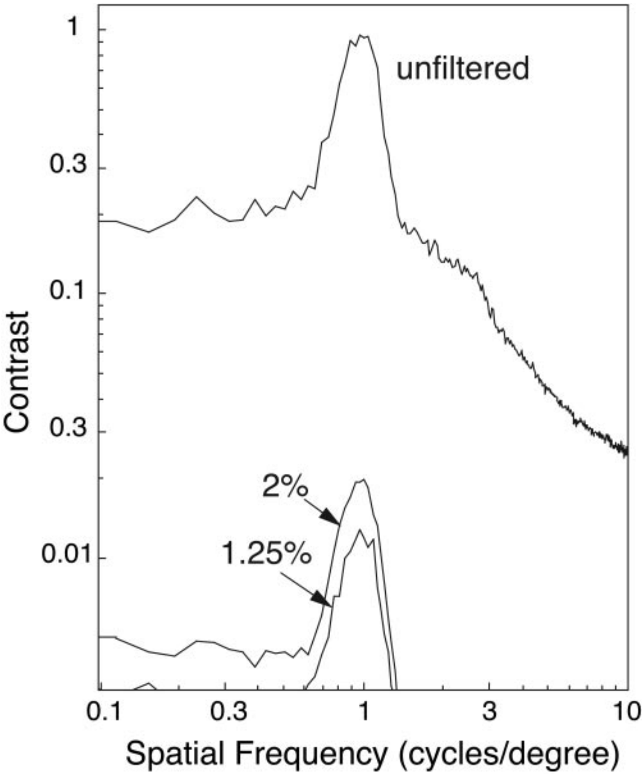

Figure 2. Fourier spatial frequency spectra for stimuli used in experiments I, II, and IV.

Between stimulus presentations, a stimulus consisting of two "flat" distracter stimuli was presented. No fixation point was used, but the stimuli were large enough and at this age the fovea is immature enough that we presumed that visual performance was largely mediated via the extrafoveal retina.[16]

Experimental Design. The data were collected by the method of constant stimuli. Each infant was tested for only one session at a given age in weeks. Each session for each infant tested two horizontal binocular disparities, each presented in crossed and uncrossed disparity configurations. In most cases, the 17.5 minutes were run in the same sessions with the 70-minute disparity stimuli, and the 35 minutes were run with the 105-minute disparity stimuli. The 7 minutes were generally run with the 17.5-minute disparity stimuli in the case of younger infants and with the 3.5-minute disparity stimuli in the case of older infants. Each block of trials contained eight stereopsis trials: two examples of each of two disparities: crossed and uncrossed. Also, each stimulus set of approximately 10 blocks included about three "easy" stimuli. An easy stimulus was a patch of the stereogram texture presented on one side to the right eye, the left eye, or both eyes, whereas the rest of the stimulus (including the test field in the opposite eye in the case of monocular textures) was at the mean luminance (see Fig. 1B for a monocular example). The purpose of these "easy" stimuli was to show the tester an occasional example of easy-to-see looking behavior. We retained all data in excess of 10 trials per stimulus per age (mean = 20 ± 5.8 trials), which amounted to 303 of the 307 sessions' worth of data we collected. Each data point was the average of data collected on approximately six infants (mean, 6.4 ± 1.9).

Procedure. The data were collected using forced-choice preferential looking (FPL),[17] modified as described in Brown and Miracle.[18] The adult tester held the infant in a infant sling facing the stimulus array at a testing distance of 0.9 m. The apparatus prevented the adult tester from seeing the test stimuli. The infant wore the analyzer spectacles described earlier, which directed the stimuli to the right and left eyes. The apparatus included an infrared light source, located just below the stimulus display, and an inconspicuous hot mirror suspended in front of the stereoscopically "flat" part of the top edge of the stimulus. An infrared-sensitive video camera was trained on the infant's face, via the hot mirror and through the polarizing lenses and sent its signal to a video monitor, which displayed an image of the infant's eyes and allowed the tester to observe the infant's looking behavior. In each trial, the tester judged whether the test stimulus was on the right or the left side of the screen, based on the infants looking behavior. An experimenter tabulated the results of each trial and advanced the presentation of the stimulus slides (PowerPoint; Microsoft). The duration of a typical trial was less than 10 seconds.

Data Analysis. The data were the percentage of trials on which the tester judged correctly the left–right location of the test stimulus. The data were analyzed in two ways: first by nonlinear regression (SPSS, Chicago, IL). This allowed us to fit a descriptive model to the data, which in turn allowed us to estimate the maximum level of performance at each age. Second, we defined a criterion level of performance, the lowest level of performance that indicated stereopsis. We chose 56.3% correct as our criterion, because that is the minimum level of performance that would be significantly above 50% correct for five infants tested for 20 trials each, one-tailed, assuming binomial variability in performance. We obtained stereoacuity and d-max data from the group data by using linear interpolation to this criterion.

Results

The average data are shown as psychometric functions in Figure 3. The data were best fit by the descriptive model:

$$C' = \left[1 + \log_{10}\left(\frac{10^{(A-12)^{0.5}}}{(10^{13.78^{.5}} + 10^{(A-12)^{0.5}})}\right)\right] \cdot (-0.136 + 0.034D - 0.00063D^2 + 3.11 \cdot 10^{-6}D^3) - 0.006 \cdot P \cdot (A - 12) \quad (1)$$

where

$$C' = 2(C - 0.5)$$

C is the fraction correct performance. The first factor in the first term in equation 1 describes the effects of A, infant age in weeks: sudden improvement at first, little improvement after about age 14 weeks. The second factor in the first term described the dependence on D, the binocular disparity in minutes of VA. The function was unimodal, and its maximum was at 35 arc min of disparity. In the second term, P was $+1$ for crossed disparity and -1 for uncrossed disparity; the statistical significance of the interaction between age and disparity indicated that crossed and uncrossed disparity were significantly different in older infants.[19] We tried other models, including polynomial effects of A; higher-order polynomial effects of D; main effects for A, D, and P; and other interaction terms between A, D, and P, but this model was the simplest for which the confidence intervals around all the constants excluded zero. The model accounted for 19.4% of the variance in the whole data set.

Figure 4 shows three infant psychometric functions and the fits of equation 1. The data at ages 12 and 13 weeks were near chance, and, not surprisingly, did not fit the model (or any other model) very well (Figs. 4A, 4B). In contrast, the performance of the older infants (aged 14–20 weeks) was better, and the data were generally quite similar to one another (Fig. 3). We averaged those data (Fig. 4C) and compared the average to the model predictions for 20 weeks. The average data and the model were in good agreement.

The stereoacuities obtained from linear interpolation to the critical level of percent correct appear in Figure 5A (diamonds), along with the predicted values from equation 1 (the bottom curve) and the crossed stereoacuity data of Birch and

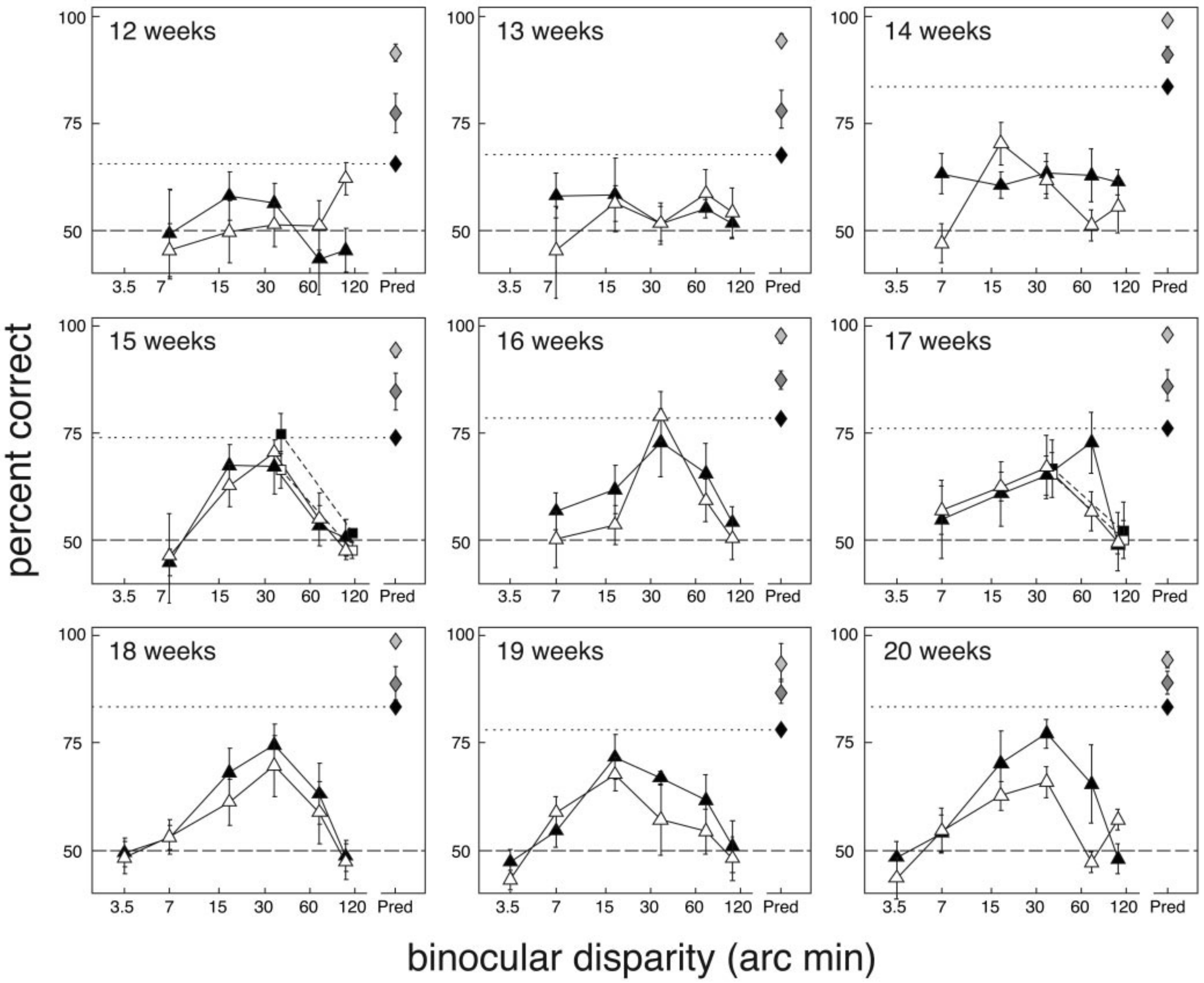

Figure 3. Psychometric functions ± SEM for stereopsis in infants aged 12 to 20 weeks. *Dashed lines*: chance performance. *Triangles*: distracter was an equal amount of vertical disparity; *squares*: distracter was zero disparity; *black triangles* and *squares*: crossed disparity; *white triangles* and *squares*: uncrossed disparity; *dotted lines* and *black diamonds* above "Pred": predicted maximum level of performance, as estimated from monocular performance in detecting the stereogram texture (*medium-gray diamonds*). Binocular detection performance (*light gray diamonds*) is shown for comparison.

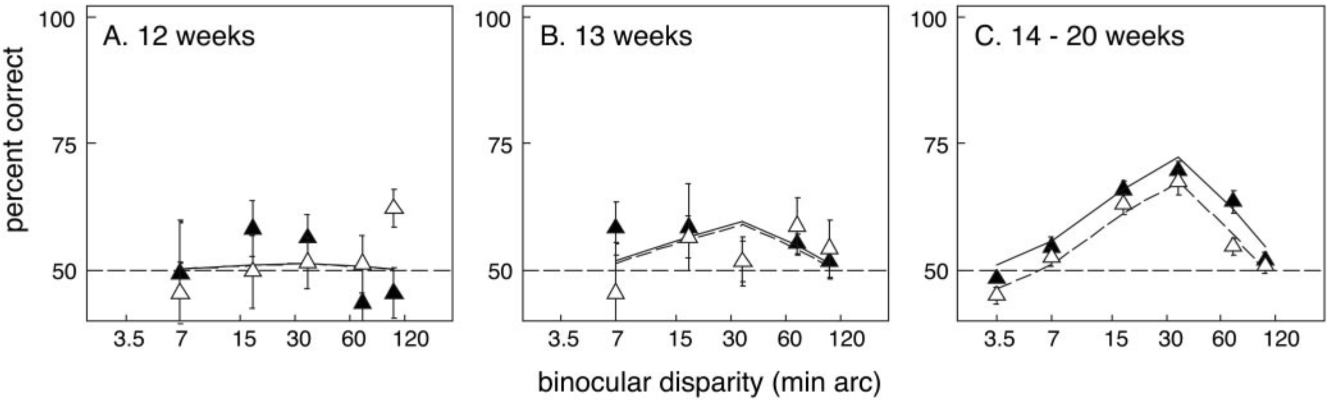

Figure 4. Fits of the model (equation 1). *Solid curves*: crossed disparity; *dashed curves*: uncrossed disparity. *Horizontal dashed line*: chance. (**A**, **B**) Data from Figure 3 and the corresponding model fits; (**C**) average data for ages 14 to 20 weeks; SEM bars are from the age data; model fits are for 20 weeks.

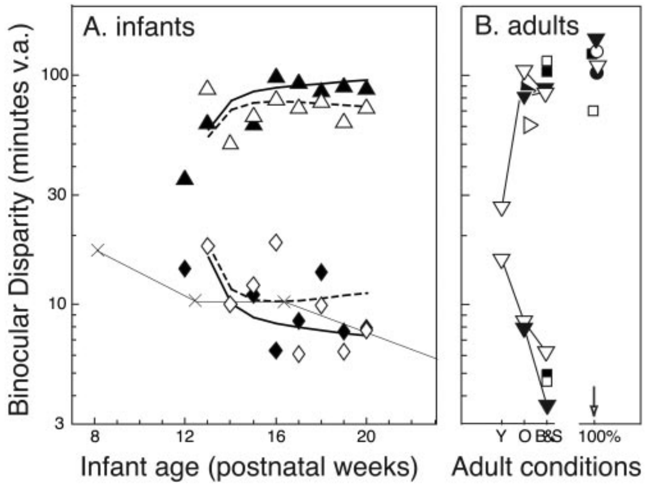

FIGURE 5. Stereoacuity and d-max on infants and adults. *Black symbols*: crossed disparity; *white symbols*: uncrossed disparity. *Top symbols* (above 25 minutes): d-max data; *bottom symbols*: stereoacuity data. (**A**) Infant data from experiment I. *Triangles*: d-max; *diamonds*: stereoacuity; *bold lines*: equation 1 (*solid*: crossed disparity; *dashed*: uncrossed disparity). ×, data from Birch and Salomao.[20] (**B**) Adult data run at 100% contrast and under contrast conditions intended to mimic the visibility of the stereogram texture to younger (Y) and older (O) infants, and under the conditions of Banks and Salapatek[21] (B&S). Each symbol type represents a different adult subject; *inverted triangles*: AMB; *circles*: PNS; *squares*: AMW; *right-facing triangles*: JMT. Stereoacuity was not measured in any adult subject at 100% contrast (*arrow*).

Salomao[20] (×'s). The near-chance data on the 12- to 13-week-olds in Figures 3 and 4 were not easily interpolated to the critical stereoacuity, and so there are several missing data points at those ages in Figure 5. Stereoacuity was essentially constant at approximately 8.9 arc min between 14 and 20 weeks, in good agreement with Birch and Salomao. This was far worse than the nominal adult stereoacuity of 5 to 10 arc sec (e.g., Ref. 22 and many others).

The d-max is represented by upright triangles in Figure 5A. To our knowledge, these have not been reported in the literature before, except for a brief mention in an abstract by Wattam-Bell (*IOVS* 1995;36:ARVO Abstract 4180). Between 14 and 20 weeks, infant d-max was essentially constant at approximately 85 arc min, which was slightly below the adult d-max for random-element stereograms from the literature (e.g., Refs. 12,23). The top curve shows the d-max from equation 1.

The model (equation 1) indicated that infant performance at 35 arc min of binocular disparity was the correct estimate of the maximum percentage of correct performance, which is shown as a function of age in Figure 6.

The literature on infant stereoacuity commonly reports that only some of the infants over this age range have measurable stereopsis.[24,25] Our data at 35 arc min of disparity (Fig. 7) also show that few infants at age 12 to 13 weeks and all infants at age 20 weeks could perform the stereo task at or above our criterion level of performance. The inverted-U shape of the psychometric functions in Figure 3 indicates that when infants fail to demonstrate stereopsis, the problem is not that the equipment cannot present a large enough binocular disparity. Rather, the problem is a fundamental inability of infants to perform the stereo task. As the horizontal binocular disparities increased above approximately 35 arc min, the disparity approached the infants' d-max, and the binocular disparity became harder to detect, not easier. This was true even when the nonstimulus side of the display was "flat" and fusible (Fig. 3, 15- and 17-week-olds, squares).

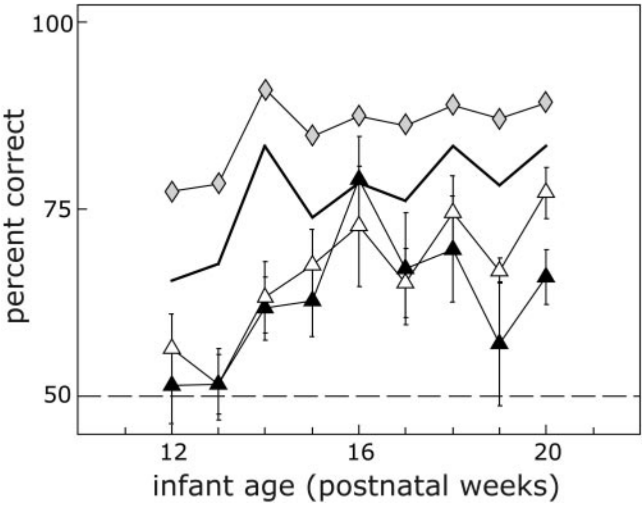

FIGURE 6. Predicted and measured maximum infant percentage of correct performance as a function of age. *Triangles*: infant data at 35 arc min of binocular disparity, from experiment I (*black*: crossed disparity; *white*: uncrossed disparity). Error bars: ± SEM of the individual infant data. *Gray diamonds*: monocular detection data from experiment III and Figure 3. *Bold line*: predicted upper boundary on infant stereo performance. Infant maximum performance was consistently 12.1% below the upper boundary, but the shape of the developmental curve was correctly predicted.

EXPERIMENT II: ADULT DATA AT 100% CONTRAST

To interpret the infant data from experiment I, we collected data on three adults by using the same stimuli as we used on infants. Clinical examination included visual acuity, cover test, ocular motility, near point convergence, vergence facility, accommodative facility, and stereopsis. We also consulted the subjects' charts at the Primary Vision Care service of The Ohio State University College of Optometry. All subjects had corrected-to-normal visual acuity in each eye, no clinically significant anisometropia, and no personal or family history of amblyopia or other serious vision disorder. Examination revealed no strabismus, and the phoria status of all subjects was within physiological normal limits. No vergence or accommodative anomalies were observed, and Randot stereoacuity was within normal limits.

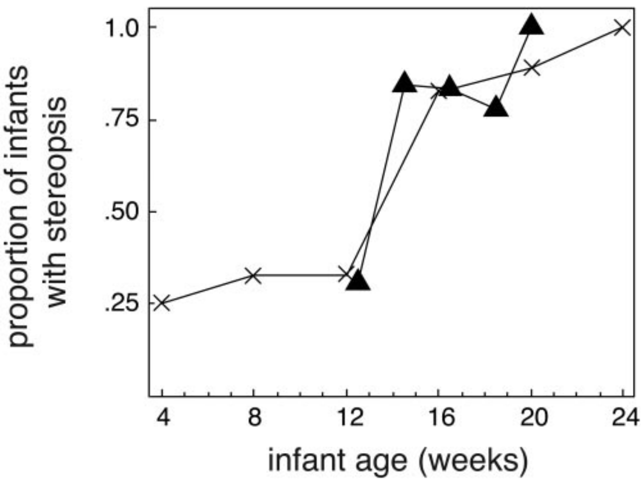

FIGURE 7. Proportion of infants with stereopsis (▲), as defined by our criterion. Data from Birch et al.[24] (×) are shown for comparison.

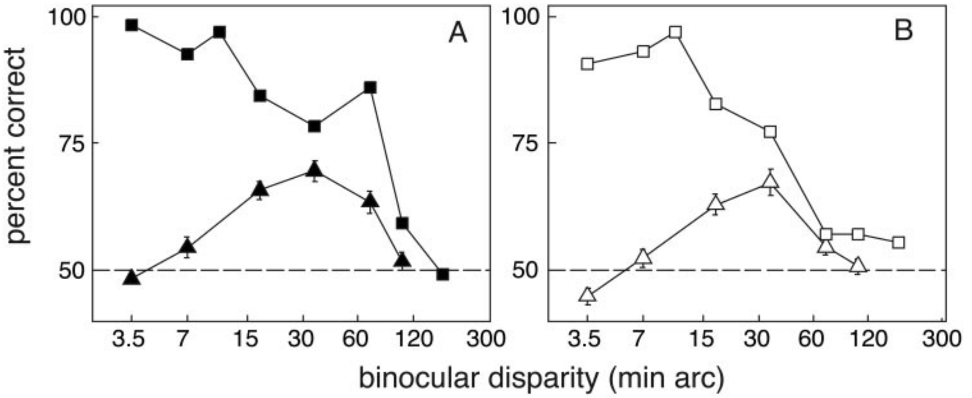

FIGURE 8. Psychometric functions on adult subject AMB (*squares*), compared to 14- to 20-week-old infant data (*triangles*), measured at 100% stimulus contrast. (**A**) Crossed disparity; (**B**) uncrossed disparity. The adult stimulus was 1.4 seconds and 6.6° from the central fixation point. Infant and adult data were similar at large binocular disparities, where d-max is measured, but were greatly divergent at small disparities, where stereoacuity is measured.

Methods

The stimuli and methods were the same as those used in the infants, with the following exceptions: (1) The stimulus was presented with a fixation point that placed the near edge of each test stimulus at 6.6° visual angle from the point of fixation. The fixation point also encouraged vergence onto the stimulus base plane and discouraged vergence onto the test or the distracter stimuli. The subject was instructed to remain fixated on the fixation point at all times. (2) Subjects could have ignored the instructions and looked at the test stimulus, but the stimulus was shown for only 1.4 seconds, which minimized this temptation. The duration was controlled by computer (PowerPoint; Microsoft) and was carefully calibrated with a photodiode. (3) The adult subjects made direct left–right judgments of the location of the test stimulus, which were recorded by the experimenter who also ran the slide presentation. (4) All binocular disparities were mixed within each block of trials by the method of constant stimuli. (5) Approximately 100 trials per data point were collected (range, 40–130 trials).

Results

At small binocular disparities, subjects reported that both the test stimulus and the distracter appeared binocularly fused, and the structure of the test stimulus was easy to see in depth. In contrast, at medium binocular disparities, the disparity was easy to see on both sides of the display. In that case, adult performance was based on whether the right- or lefthand stimulus appeared flat or rivalrous, and whether it appeared closer or farther than the base plane of the stimulus as a whole. At the largest disparities, both the test stimulus and the distracter appeared binocularly rivalrous, and the subject could not tell the difference between the test stimulus and the distracter. For clarity, we examined in detail the psychometric detection data on subject AMB, as she contributed data to all the various adult testing conditions. The stereoacuity and d-max data of all three subjects appear in Figure 5. Each of the subjects had somewhat better performance on uncrossed (AMB, AMW) or crossed (PNS) disparity, as is commonly observed.[26] The asymmetry was most pronounced in the data of AMW and least for subject PNS. All adult crossed and uncrossed data were analyzed separately.

The most obvious feature of the adult psychometric functions (Fig. 8) is that performance was good at small binocular disparities, but declined for larger disparities. Indeed, the adult data were not far from the data from the older infants above ~35 arc min of binocular disparity. This result can also be seen in Figure 5, where the adult d-max data at 100% contrast can be compared directly to the infant data in the left panel. Adult d-max at 100% contrast was generally at a slightly larger binocular disparity than in the case of infants. The other obvious feature of the adult data is that they did not decline to near chance at the smallest binocular disparities. The adult stereoacuity was therefore not measured (Fig. 5, arrow).

Discussion

To understand the implications of the similarity of d-max in infants and adults, we propose two general classes of explanation for the poor stereopsis of infants. One is a "spatial theory" and the other is a "performance theory," diagrammed in Figure 9.

The spatial theory (Fig. 9A) holds that infant stereoacuity is poor because the infant visual system extracts stereo information by comparing monocular signals over relatively longer distances than in the adult visual system. As a first approximation, this theory predicts that infant and adult psychometric functions should be similar, except that the infant data (fine curve) should be shifted to larger binocular disparities relative to the logarithmic spatial axis compared with adult data (bold curve). As infants mature, their stereo psychometric functions translate toward smaller binocular disparities (arrow), allowing

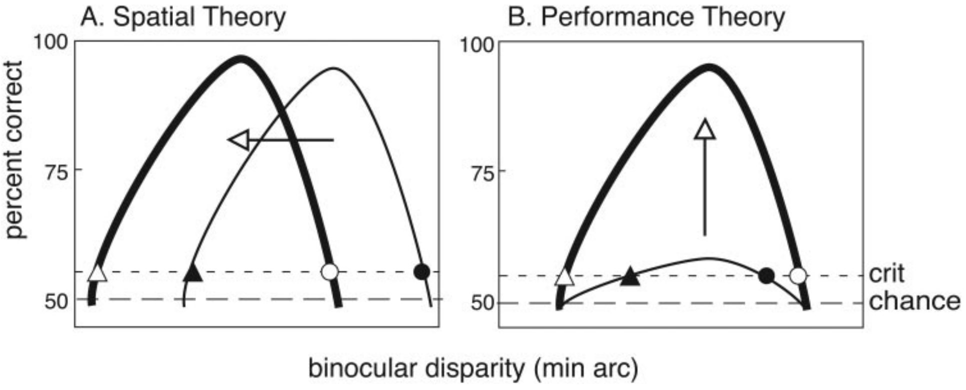

FIGURE 9. Two general theories of stereopsis development. *Curves*: psychometric functions for stereopsis as predicted by each theory; *fine lines*: infants; *bold lines*: adults; *dashed line*: chance performance level; *dotted line*: the critical value of performance for defining stereoacuity (*triangles*) and d-max (*circles*) thresholds. *Arrows*: direction of change of the psychometric function with age under each theory.

finer stereoacuity to develop (triangles). Under the spatial theory, the d-max (circles) should also become lower.

The performance theory (Fig. 9B) holds that infants differ from adults in their overall performance, but not in the spatial aspects of their binocular function. We divide performance theory into two versions. One version of performance theory says that infants do not perform the task as well as adults do, but their stereoptic capability is similar to that of adults. We will call this the "psychophysical" performance theory, because it is based on the poor overall level of infant psychophysical performance. For example, infants may not look at every stereo target they can see, they may have lapses in attention, and so forth. Psychophysical performance theory probably predicts that infant data should look like adult data, only compressed, to be closer to chance performance (Fig. 9B, fine curve). This version of the performance theory correctly predicts that infant d-max should be slightly smaller than adult d-max (Fig. 9B, black circle to the left of the white circle). However, the small-disparity end of the infant graph in Figure 8 does not look much like a compressed version of the adult data. This fact suggests that the psychophysical performance theory is not going to account for infant stereopsis much better than the spatial theory did.

A second version of the performance theory holds that infants do not perform the stereopsis task as well as adults do, because they cannot see the stereogram texture as well as adults do. We will call this the "visual" performance theory because it is based on the poor visual performance of infants in detecting the texture, and it includes more generally all differences between infants and adults that are described by the differences in their contrast sensitivity functions. Under the visual performance theory, the psychometric functions for adults may or may not look like scaled-up functions of infant psychometric functions, depending on the spatial frequency content of the stimulus.[22] In experiment III, we tested the visual performance theory to find out whether it can be disproved by what we know about infant contrast detection, or by more carefully controlled adult experiments.

The infant data from experiment I were clearly more consistent with the two performance theories than with the spatial theory, because the infant data showed a much lower maximum level of performance than did the adult psychometric data from experiment II. Furthermore, d-max moved a small amount, and generally in the direction predicted by the performance theories (Figs. 5, 8). However the adult psychometric functions from experiment II did not agree with either prediction very well, suggesting that some refinement to the theories is needed.

EXPERIMENT III: ADULT AND INFANT CONTRAST DETECTION

Our tests of the visual performance theory depended on knowing the ability of infants and adults to detect the texture stimuli used in experiments I and II. Therefore, we measured infant and adult monocular and binocular detection performance for a test stimulus consisting of a patch of the texture used in experiments I and II.

Methods

In the case of adults, the stimulus was a uniform green field at 35 cd/m^2, within which was embedded a square patch of texture of the same size as the test stimulus in the stereo experiments (Fig. 1B). Luminances and gamma corrections were calibrated with a Pritchard photometer (Photoresearch, Inc., Chatsworth, CA) in situ, through the polarizer filters and the analyzer glasses. The stimulus duration was 1.4 seconds, and the stimuli varied in Michelson contrast from trial to trial to measure psychometric functions for contrast threshold. In each trial, the stimulus could be on the right or on the left side of the central fixation point, and it could be presented to the right eye, or to the left eye, or to both eyes by means of the same arrangement of projectors and polarizers as in experiments I and II. Five very-low-contrast binocular conditions for adults were achieved by diluting the contrast from one projector with the uniform field from the other, while the subject viewed the stimulus binocularly via neutral density filters instead of polarizers. The intertrial interval contained a binocularly presented full-field stereo texture to ensure that the level of contrast adaptation was similar to that of the stereo experiments. The adult data were collected by direct psychophysical judgment of the left–right stimulus location. In the case of infants, the monocular and binocular stimuli were like those used for adults, except that the duration was as long as necessary (generally <10 seconds), and no texture was presented during the intertrial interval. Seventeen individual infants were tested, each contributing data to one or more age groups. Every infant we tested contributed at least one data set, and three sessions' data had to be discarded because there were fewer than 10 trials per contrast value. Each age group included data on approximately six infants (mean 6.1 ± 1.1). Other methods were as for experiments I and II.

Results

The monocular and binocular psychometric functions are shown for infants and adult subject AMB in Figure 10. Infants were much less sensitive to contrast than adult subject AMB (and the other adults, data not shown in Fig. 10). The adult-to-infant contrast threshold ratio, defined at 78% correct performance, was 0.027 and 0.072 for monocular detection by 12- to 13-week-olds and 14- to 20-week-olds, respectively, and similarly, for binocular detection, the ratios were 0.014 and 0.034 for the younger and older infants, respectively. Furthermore, the infant psychometric functions had a much shallower slope than the adult functions, and infant performance did not reach 100% correct, even at 100% contrast. The monocular and binocular detection performance obtained using the 100% contrast texture appear as the light- and medium-gray diamonds in Figure 3, and monocular detection performance at 100% is plotted as a function of age in Figure 6.

Given the drastic differences between infant and adult performance in detecting the stereogram texture (Fig. 10), it does not seem quite appropriate to compare the results of experiments I and II, as we did in Figures 4 and 8. We pursued this idea quantitatively in experiment IV.

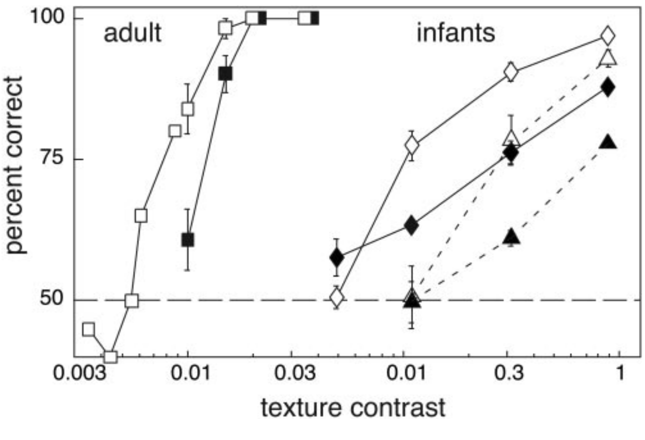

FIGURE 10. Psychometric functions of adult subject AMB (*squares*) and 12- to 13-week-old (*triangles*) and 14- to 20-week-old (*diamonds*) infants for detecting the stereogram texture presented monocularly (*black symbols*) and binocularly (*white symbols*). The five lowest binocular data points were collected separately.

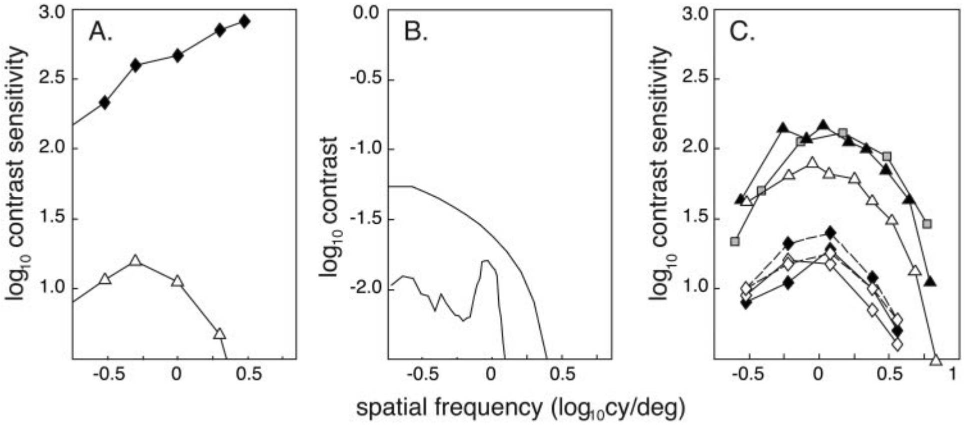

FIGURE 11. Infant and adult contrast sensitivity data from the literature. (**A**) Banks and Salapatek[21] (B&S): 3-month-old infant and adult data collected during free fixation. (**B**) *Top curve*: the difference between the infant and adult psychometric functions in (**A**) was the B&S filter used to create the stimuli for Figures 12C and 12F. *Bottom curve*: Fourier spectrum of a typical B&S-filtered stimulus. (**C**) Adult data collected in the visual periphery: Kelly[34] (*squares*) and Rovamo and Virsu[33] (*black triangles*: temporal visual field; *white triangles*: nasal visual field); infant data collected during free fixation on 3- (*solid curves*) and 6-month-olds (*dashed curves*; *white triangles*: monocular; *black triangles*: binocular). Data are from Drover et al.[32]

EXPERIMENT IV: ADULT EXPERIMENTS AT LOW CONTRAST

One way of testing the visual performance hypothesis of infant stereopsis is to test infant and adult stereopsis with equally visible stereogram textures. If the poor performance of infants is due to their inability to see the stereogram texture, then this manipulation should bring infant and adult stereopsis data into agreement. We know of no way of increasing the visibility of our texture to infants, because the texture is already at 100% contrast and its spatial frequency maximum is near the maximum of the classic infant contrast sensitivity function. However, it is easy to reduce the visibility of the texture to adults. It is well known from the literature that adult stereo performance is worse at low contrast than at high contrast.[27–31] This experiment addressed the question of whether the infant and adult data agree quantitatively when the stereogram textures are equally visible to infant and adult subjects.

Methods

In the present experiment, we created stereogram textures that were filtered in the spatial domain until they were as visible to adults as to infants. We want to stress that we are not implying that the low-contrast textures looked the same to adults as 100% contrast gratings did to infants. That would have required that we mimic the infant's introspective personal visual experience, which we plainly cannot do. Instead, we tried three contrast manipulations, each designed to make the texture as hard to see for our adult subject as the 100% contrast stimulus was for infants. Aside from the manipulated contrast, the methods of experiment IV were the same as in experiment II.

In our first effort, we reduced the contrast at each spatial frequency by a filter defined as the ratio of the infant and adult contrast sensitivity functions (CSFs) shown in Banks and Salapatek.[22] This is not the ideal way of doing the experiment, because the adults in that experiment were tested in the fovea. However, this manipulation has the advantage that it is based on classic contrast sensitivity functions on infants and adults, which were measured in the same experiment. The CSFs from Banks and Salapatek appear in Figure 11A. The maximum transmission of the filter was 0.035, and it tapered off at higher spatial frequencies (shown against logarithmic axes in Fig. 11B, top curve). The contrast spectrum of the filtered stimulus also appears in Figure 11B (bottom curve).

In our second effort, we based our contrast filter on our data from Experiment III. We began by comparing the CSFs of infants to adult CSF data from the literature. We chose the infant data in Drover et al.[32] which included monocular and binocular CSFs in 3- and 6-month-olds. We chose two adult studies in which the test stimulus was presented at about the same eccentricity at which we measured the stereo performance for our adults (Rovamo and Virsu,[33] 7.5°; Kelly,[34] 8°). Those CSFs appear in Figure 11C. The remarkable result of this comparison was that the infant CSFs were approximately the same shape as the adult CSFs, except for being shifted down to lower $\log_{10}$ sensitivities. This suggested that we could approximate the infant's detectability of the stereogram texture by simply scaling the contrast to a lower value. A comparison of infant and adult psychometric functions in Figure 10 indicates that adult subject AMB's performance reached 100% correct only when the stereo texture was at a contrast of 0.02. Her performance was estimated by interpolation to be near the maximum performance of the younger infants when the stereo texture was approximately 0.0125 contrast. Therefore, we collected data on subject AMB using stimuli at those contrasts. Psychometric data (not shown) on subject JMT indicated that her performance reached 100% between the contrasts of 0.01 and 0.015, and she was tested using the 0.125 contrast test stimulus, to approximate the texture visibility to the older infants.

Results

The psychometric functions of subject AMB appear in Figure 12. The agreement between the infant and the adult data was qualitatively good. When AMB was tested at a contrast of 0.0125 (Figs. 12A, 12D), her near-chance performance resembled the near-chance performance of the youngest infants. Her maximum performance was low, and stereoacuity and d-max were measured only for uncrossed disparity (Fig. 5, right, Y). In contrast, when AMB was tested at a contrast of 0.02, her performance was better, and her data resembled the data of the older infants, both in terms of her maximum performance (Figs. 12B, 12E), and in terms of her stereoacuity and d-max (Fig. 5, right, O). The results when the stimuli were adjusted using the low-pass filter from Figure 11B resembled the results in older infants (Figs. 12C, 12F; Fig. 5, B&S). The psychometric functions of the other two subjects (not shown) were similar to those of AMB. D-max data and (when available) stereoacuity data from the other adult subjects appear in Fig. 5.

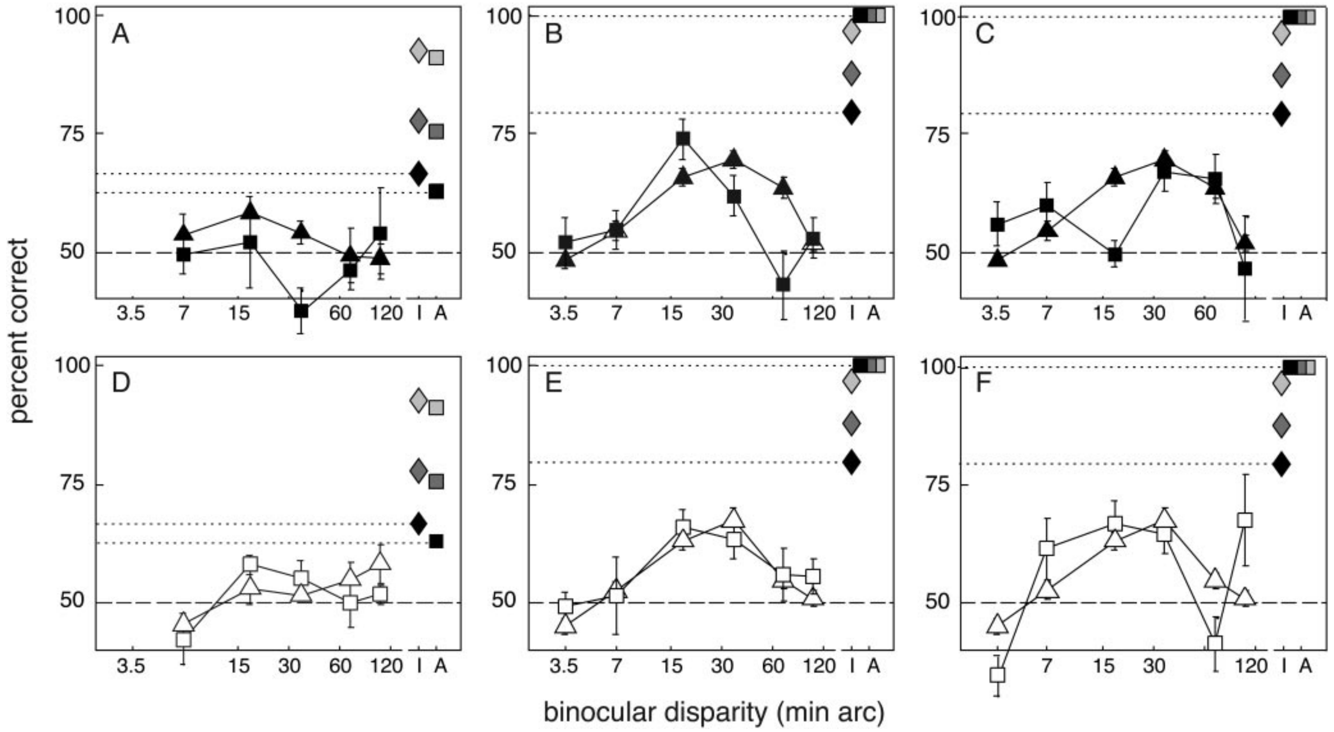

FIGURE 12. Adult psychometric functions (*squares*: subject AMB), collected under conditions designed to mimic the visibility of the stereograms to infants. (**A–C**) Crossed disparity; (**D–F**) uncrossed disparity. (**A**, **D**) At 1.25% contrast, compared with the average of the infant data at 12 and 13 weeks (*triangles*); (**B**, **E**) at 2% contrast, compared with the 14- to 20-week data averaged from Figure 3 (*triangles*). (**C**, **F**) 3.5% low-pass filter, designed from the contrast sensitivity data of Banks & Salapatek[21] (Figs. 11A, 11B), compared with the 14- to 20-week-old data. I and A, *light* and *medium gray* symbols: binocular and monocular, infant and adult contrast detection data, from experiment III; *black symbols* and *dotted lines*: predicted upper boundary performance. The overall agreement between the infant and adult data suggests that infant stereo performance is limited by the visibility of the stereogram texture.

When comparing the psychometric results shown in Figures 8 and 12, it is remarkable that when the stimuli were reduced in contrast, the adult lost much more sensitivity to small disparities than to large disparities, and stereoacuity changed from excellent (lower than our equipment could measure) to poor (higher than 3.5 arc min). This qualitative change in adult performance was probably due to the loss of the high-spatial-frequency shoulder and tail of the Fourier energy spectrum (Fig. 2), which corresponded to the sharp edges of the texture elements. This high-spatial-frequency energy was highly visible to adults, and it probably mediated the superior adult performance for small binocular disparities at 100% contrast (Fig. 8).[22] In contrast to its dramatic effect on stereoacuity, low-pass filtering is known to have little effect on d-max in adults.[12,22] Therefore, it is not surprising that the loss of high spatial frequencies due to the reduced contrast of our stimuli had little effect on adult d-max.

We believe that these changes in adult performance occurred precisely because the contrast manipulations succeeded in mimicking the visibility of the stereograms to infants. Like adults tested at low contrast, infants could never see the high-spatial-frequency components of our stereo stimuli because of their poor contrast sensitivity and poor resolution acuity (Figs. 11A, 11C). Thus, the lack of visible high-spatial-frequency components in the stimuli, for infants (tested at 100% contrast) and for adults (tested at very low contrast), can explain the pattern of results illustrated in Figures 5 and 12: poor stereoacuity but good d-max in both groups of subjects. This pattern of results suggests that the critical immaturity that limits infant stereo performance is indeed the inability of infants to see the stereogram texture very well.

GENERAL DISCUSSION

The infant data reported in Experiment I indicated that infant stereoacuity and maximum stereo performance are immature over the age range between postnatal ages 12 and 20 weeks. Whereas a typical adult stereoacuity is approximately 10 arc sec of binocular disparity, infant stereoacuity improved only to 8.9 arc min (averaged over the ages 14–20 weeks). The surface portrayed by the 100% contrast stereogram at 35 arc min of binocular disparity was solidly fused by the adult into a single percept, and appeared clearly at a different depth from its surrounding base plane to the normal adult observer, yet infants did not detect it above 79% correct (70% correct, averaged over ages 14–20 weeks; Fig. 4C). In contrast to these immature features of infant stereoptic psychometric functions, the value of d-max, the largest value of binocular disparity that was distinguished from binocularly rivalrous vertical binocular disparity, was slightly lower in infants than in the adults. This similarity of d-max in adults and infants rules out an explanation of infant stereopsis based on extra long-distance lateral interactions in infant stereo vision, which may preclude fine stereoacuity.

Experiment IV was designed to test the hypothesis that infant stereopsis performance was poor because infants could not see the stereogram texture well enough to do the stereo task. In that experiment, the contrast of the stereogram texture was filtered so that it was as visible to the adult as the 100% contrast stereogram texture was to infants. Consistent with the visual performance hypothesis, adult stereo performance was close to infant stereo performance under these conditions.

Infant Monocular Detection Performance

In this discussion, we will use the infant monocular detection data from experiment III to predict infant stereopsis performance under the visual performance hypothesis. The key observation is that no subject (infant or adult) can perceive the depth of a stereogram unless the stereogram texture is simultaneously visible to the two eyes. Therefore, an upper boundary on infant stereo performance can be estimated from the infant monocular detection data shown as the medium-gray diamonds in Figures 3 and 6. We compared these estimates to the maximum performance of infants. This analysis will provide evidence about the limitations of infant stereopsis that does not depend on comparing infants with adults.

The first step in this analysis was to estimate the probability of seeing the texture monocularly via the right and left eyes, respectively ($P_S[L]$ and $P_S[R]$), from the proportion of correct trials on the monocular forced-choice detection task ($P_C[L]$ and $P_C[R]$), using the standard correction for guessing, for example:

$$P_S[L] = 2 \cdot (P_C[L] - 0.5) \qquad (2)$$

Then, the probability of seeing the stereogram texture simultaneously via both the right and left eyes ($P_S[B]$) was obtained by multiplying the probability of seeing the texture via the right eye and the left eye, $P_S[L]$ and $P_S[R]$.

$$P_S[B] = P_S[L] \cdot P_S[R] \qquad (3)$$

Finally, we used the inverse of equation 2 to predict the predicted maximum possible stereopsis performance of infants and adults from equation 3.

The results of this calculation are summarized in Figure 6. The gray diamonds are the monocular detection data from Figure 3, and the bold line in Figure 6 is the predicted upper boundary on stereopsis performance. The upper boundary also appears as the black diamonds and horizontal dotted lines in Figures 3 and 12. The black and white triangles in Figure 6 are the values of stereo performance at 35 arc min, which our model (equation 1) indicated to be the best estimate of infant maximum performance at each age. Figure 6 shows that the stereopsis data fell below the upper boundary of performance, as predicted. Remarkably, the average difference between the observed performance and predicted upper boundary was only 12.1%. Furthermore, a two-way linear ANOVA (SPSS) on the differences between predicted upper boundary and observed performance revealed no change with age ($P = 0.061$), no difference based on crossed–uncrossed polarity ($P = 0.693$), and no interaction between factors for age and polarity ($P = 0.744$), indicating that the upper boundary model correctly predicted the shape of the curve of development.

Adult Monocular Detection Performance

It is not clear that the remaining difference between the predicted upper boundary and the observed maximum of the psychometric functions in Figure 3 was in fact the result of any critical immaturity. We used the contrast detection data for adult subject AMB to predict the upper boundary of her psychometric functions for stereopsis (Fig. 12, dotted lines–black squares). For the conditions mimicking the data of the youngest infants, the adult data showed no better agreement with the predicted upper boundary than the infant data did, and in the condition mimicking the older infants, no adult errors were predicted, and so the adult data fell far below the upper boundary. Subjectively, subjects found the stereo task at low contrast to be "unpleasantly challenging." Apparently, the stereogram texture needs to be above its contrast-detection threshold for infants and adults to perform the stereo task. No critical immaturity is required to explain the difference between the predicted and observed maximum stereo performance of infants in Figure 6.

CONCLUSIONS

The results of these experiments indicate that infant stereopsis performance was far below adult stereopsis performance when both were tested with the same 100% contrast stereogram stimulus. However, this poor level of infant performance is probably not due to a special critical immaturity that limits infant stereopsis but not other aspects of infant visual function. Instead, our experiments suggest that infant stereopsis is poor because of their overall poor visual performance, particularly their poor performance in detecting the stereogram texture. We supported this conclusion by two independent lines of evidence. First, adults performed as badly as infants did, when adults were tested in the extrafoveal visual field, using a low-contrast stereogram texture that mimicked the visibility of the texture to infant observers. Second, infant stereopsis performance and adult stereopsis performance were close to, but did not reach, the optimum level of stereopsis performance predicted from their own monocular detection performance of the stereogram texture.

Where does this leave us with our original reasons for studying infant binocular stereopsis? Those who study stereopsis because it is critically limited by binocular interactions in the visual cortex will probably be disappointed to hear that it is not likely to be possible to study those interactions psychophysically in infants. We do not know whether infants would have better stereopsis "if only" they could see the stereogram texture better. However, we suspect that this is not a question that can be answered easily, because the stereogram texture in this study was close to the most visible stimulus according to the infant contrast sensitivity function, which is dominated by extrafoveal mechanisms. Perhaps future research that investigates the maturation of the infant fovea will shed light on this important problem. In contrast, clinicians who use stereopsis testing as a way of discovering whether a particular infant patient can see out of both eyes should be heartened. Our results suggest that the critical immaturity that limits infant stereo performance is indeed the very monocular visual performance that the clinician needs to measure.

Acknowledgments

The authors thank The Ohio State University College of Optometry for providing the infant eye examinations, and Angela M. White, Lan Ta, Ebla Dahir, and Jackie M. Taylor for assistance with data collection.

References

1. Shea SL, Fox R, Aslin RN, Dumais ST. Assessment of stereopsis in human infants. *Invest Ophthalmol Vis Sci*. 1980;19:1400–1404.
2. Held R, Birch E, Gwiazda J. Stereoacuity of human infants. *Proc Natl Acad Sci USA*. 1980;77:5572–5574.
3. Petrig B, Julesz B, Kropfl W, Baumgartner G, Anliker M. Development of stereopsis and cortical binocularity in human infants: electrophysiological evidence. *Science*. 1982;213:1402–1405.
4. Read J. Early computational processing in binocular vision and depth perception. *Prog Biophys Mol Biol*. 2005;87:77–108.
5. Simons K. Preschool vision screening: Rationale, methodology and outcome. *Surv Ophthalmol*. 1996;41:3–30.
6. Held R. Two stages in the development of binocular vision and eye alignment. In: Simons K, ed. *Early Visual Development Normal and Abnormal.* New York: Oxford University Press; 1993.

7. Schor CM. Development of stereopsis depends upon contrast sensitivity and spatial tuning. *J Am Optometric Assoc*. 1985;56: 628–635.
8. Chino YM, Smith EL III, Hatta S, Cheng H. Postnatal development of binocular disparity sensitivity in neurons of the primate visual cortex. *J Neurosci*. 1997;17:296–307.
9. Farell B. Two-dimensional matches from one-dimensional stimulus components in human stereopsis. *Nature*. 1998;395:689–693.
10. Patel SS, Ukwade MT, Stevenson SB, Bedell HE, Sampath V, Ogmen H. Stereoscopic depth perception from oblique phase disparities. *Vision Res*. 2003;43:2479–2492.
11. Patel SS, Bedell HE, Sampat P. Pooling signals from vertically and non-vertically orientation-tuned disparity mechanisms in human stereopsis. *Vision Res*. 2006;46:1–13.
12. Wilcox LM, Hess RF. Dmax for stereopsis depends on size, not spatial frequency content. *Vision Res*. 1995;35:1061–1069.
13. Birch EE, Gwiazda J, Held R. The development of vergence does not account for the onset of stereopsis. *Perception*. 1983;12:331–336.
14. Braddick O. Binocularity in infancy. *Eye.* 1996;10:182–188.
15. Hainline L, Riddell PM. Binocular alignment and vergence in early infancy. *Vision Res*. 1995;35:3229–3236.
16. Brown AM, DobsonV, Maier JS. Infants' visual acuity at scotopic, mesopic, and photopic luminances. *Vision Res*. 1987;27:1845–1858.
17. Teller DY. The forced-choice preferential looking procedure: a psychophysical technique for use with human infants. *Infant Behav Dev*. 1979;2:135–53.
18. Brown AM, Miracle JA. Early binocular vision in human infants: limitations on the generality of the Superposition Hypothesis. *Vision Res*. 2003;43:1563–1574.
19. Birch EE, Gwiazda J, Held R. Stereoacuity development for crossed and uncrossed disparities in human infants. *Vision Res*. 1982;22: 507–513.
20. Birch EE, Salomao S. Infant random dot stereoacuity cards. *J Pediatr Ophthalmol Strabismus*. 1995;35:86–90.
21. Banks M, Salapatek P. Acuity and contrast sensitivity in 1-, 2-, and 3-month-old human infants. *Invest Ophthalmol Vis Sci*. 1978;17: 361–365.
22. Hess RF, Chang HL, Wang Y-Z. Luminance spatial scale and local stereo-sensitivity. *Vision Res*. 2002;42:331–342.
23. Glennerster A. Dmax for stereopsis and motion in random dot displays. *Vision Res*. 1998;38:925–35.
24. Teller DY. Scotopic vision, color vision, and stereopsis in infants. *Curr Eye Res*. 1983;2:199–210.
25. Birch EE, Shimojo S, Held R. Preferential-looking assessment of fusion and stereopsis in infants aged 1–6 months. *Invest Ophthalmol Vis Sci*. 1985;26:366–370.
26. Richards W. Anomalous stereoscopic depth perception. *J Opt Soc Am*. 1971;31:410–414.
27. Westheimer G, McKee SP. Stereogram design for testing local stereopsis. *Invest Ophthalmol Vis Sci*. 1980;19:802–809.
28. Halpern DL, Blake RR. How contrast affects stereoacuity. *Perception*. 1988;17:483–495.
29. Legge GE, Gu YC. Stereopsis and contrast. *Vision Res*. 1989;29: 989–1004.
30. Schor C, Heckmann T. Interocular differences in contrast and spatial-frequency: effects on stereopsis and fusion. *Vision Res*. 1989;29:837–847.
31. Stevenson SB, Cormack LK. A contrast paradox in stereopsis, motion detection, and vernier acuity. *Vision Res*. 2000;40:2881–2884.
32. Drover JR, Earle AE, Courage, ML, Adams RJ. Improving the effectiveness of the infant contrast sensitivity card procedure. *Optom Vision Sci*. 2002;79:52–59.
33. Rovamo J, Virsu V. An estimation and application of the human cortical magnification factor. *Exp Brain Res*. 1979;37:496–510.
34. Kelly DH. Retinal inhomogeneity. I. Spatiotemporal contrast sensitivity. *J Opt Soc Am A*. 1984;1:107–113.