Visual Acuity, Crowding, and Stereo-Vision Are Linked in Children with and without Amblyopia

John A. Greenwood,[1–4] *Vijay K. Tailor,*[2,5] *John J. Sloper,*[2,5] *Anita J. Simmers,*[6] *Peter J. Bex,*[7] *and Steven C. Dakin*[1,2]

Purpose. During development, the presence of strabismus and anisometropia frequently leads to amblyopia, a visual disorder characterized by interocular acuity differences. Although additional deficits in contrast sensitivity, crowding (the impaired recognition of closely spaced objects), and stereoacuity are common, the relationship between these abilities is unclear.

Methods. We measured the covariation between these four abilities in children 4 to 9 years of age ($n = 72$) with strabismus, anisometropia, or mixed strabismus/anisometropia, and unaffected controls. Children reported the orientation of a target (a modified "Pac-Man," similar to Landolt-C stimuli) using four "ghosts" as references. Using a modified staircase procedure we measured threshold size (acuity), contrast detection, foveal crowding (the minimum separation between target and ghost-flankers supporting accurate identification), and stereoacuity (with random-dot stereogram ghosts).

Results. Group averages revealed significant interocular differences (IODs) in acuity for all three clinical groups (0.2–0.3 log minutes), and significant crowding IODs for the strabismic and mixed groups (0.6 and 0.4°, respectively). Nonetheless, crowding IODs were correlated with acuity IODs in all four groups (r values between 0.43 and 0.59 and $P < 0.05$; $P = 0.07$ in the mixed group). Similarly, the occurrence of stereo-blindness (most common in strabismic and mixed groups) was associated with a significant increase in IODs for both acuity and crowding (each $P < 0.05$). No correlations were found with contrast detection.

Conclusions. Our results demonstrate an association between IODs in acuity and crowding and, furthermore, between these IODs and the presence of stereo-vision. We suggest that the deficits derived from strabismus and anisometropia lay along a continuum with abilities observed during normal development. (*Invest Ophthalmol Vis Sci.* 2012;53:7655–7665) DOI: 10.1167/iovs.12-10313

It is well known that abnormal visual experience can have a profound effect on the development of the visual system. In the case of children with anisometropia (unequal refractive error) and/or strabismus (ocular misalignment), a common outcome is unilateral amblyopia, a developmental visual disorder whose defining feature is reduced acuity in one eye that persists despite optical correction.[1,2] The development of amblyopia is also associated with a range of additional deficits, including anomalies in contrast detection,[2–6] elevated visual crowding,[7–9] and reduced stereoacuity or complete stereo-blindness.[2,10] Although each is becoming increasingly well understood in isolation, little is known about the relationship between these deficits, particularly during childhood when strabismus and anisometropia first develop.

Aside from the definitive acuity deficits, amblyopia is most frequently associated with deficits in stereo-vision. This can range from an elevation in stereoacuity thresholds, the magnitude of which correlates with the interocular difference in acuity,[10–12] to complete stereo-blindness, which is particularly common in strabismus[2,13,14] and cases of severe anisometropia.[2,10,12,15] The link between these deficits in acuity and stereo-vision has been attributed to altered neural selectivities within cortical area V1: the abnormal visual experience that causes amblyopia may shift both the selectivity of monocular neurons away from the amblyopic eye and reduce the number of binocular neurons.[1,2,16–18] Improvements in visual acuity resulting from binocular therapy[19,20] further suggest that there may be a causal link between acuity and stereo-vision, although a similar training regime focused on stereoacuity did not improve acuity thresholds.[21] The relationship between these impairments in stereo-vision and other amblyopic deficits is unknown.

A more complex pattern of deficits has been observed for achromatic contrast sensitivity. Although significant impairments in contrast sensitivity have been observed in the amblyopic eye of adults,[4–6] amblyopic children may show milder deficits.[3] Additionally, although the amblyopic eye of anisometropes typically supports higher contrast-detection thresholds than does the fellow eye, in strabismus there have been reports of *improved* contrast-detection thresholds when observers use their amblyopic eye.[2,6] It is possible that these changes reflect a redeployment of binocular cells as monocular cells,[2] which may suggest a relationship between these anomalies in contrast detection and the stereo-deficits described above. The variation in contrast-detection thresholds

From the [1]UCL Institute of Ophthalmology, University College London, London, United Kingdom; the [2]National Institute for Health Research Biomedical Research Centre for Ophthalmology at Moorfields Eye Hospital National Health Service Foundation Trust, London, United Kingdom; [3]Laboratoire Psychologie de la Perception, Université Paris Descartes, Sorbonne Paris Cité, Paris, France; [4]Centre National de la Recherche Scientifique, UMR 8158, Paris, France; [5]Strabismus and Paediatric Service, Moorfields Eye Hospital, London, United Kingdom; [6]Vision Sciences, Department of Life Sciences, Glasgow Caledonian University, Glasgow, United Kingdom; and the [7]Schepens Eye Research Institute, Harvard Medical School, Boston, Massachusetts.

Supported by the Special Trustees of Moorfields, the UK Department of Health through the award made by the National Institute for Health Research to Moorfields Eye Hospital and UCL Institute of Ophthalmology for a Biomedical Research Centre for Ophthalmology, and a Marie Curie Intra-European Fellowship (JAG).

Submitted for publication June 1, 2012; revised August 14 and September 29, 2012; accepted October 5, 2012.

Disclosure: **J.A. Greenwood**, None; **V.K. Tailor**, None; **J.J. Sloper**, None; **A.J. Simmers**, None; **P.J. Bex**, None; **S.C. Dakin**, None

Corresponding author: John A. Greenwood, Laboratoire Psychologie de la Perception, Université Paris Descartes, Paris, France; john.greenwood@parisdescartes.fr.

also suggests that the nature of these effects depends on both the progression and the type of amblyopia.

Finally, a particularly striking functional consequence of amblyopia is an elevation in *crowding*, the impairment in object recognition that occurs when an otherwise visible target object is flanked by similar elements.[7,9,22] In "normal" adult vision, crowding is widely viewed as a "compulsory pooling" of target and flanker signals[23–25] that is strongest in the peripheral field and minimal in the fovea.[9,26] In amblyopia, particularly when accompanied by strabismus, crowding is clearly evident in the fovea and can extend over spatial extents similar to crowding in the typical periphery.[7–9] Elevated foveal crowding has also been observed in typically developing children,[27–29] which may persist as late as 11 years.[29] Indeed, linear (so-called *crowded*) letter charts are generally considered better detectors of amblyopia in children than isolated letters.[28,30,31] The relationship between these three instances of crowding—in the "normal" periphery, the developing fovea, and the amblyopic fovea—is currently unclear. Many theories have been proposed to account for crowding,[9] with one view describing both peripheral and amblyopic forms as the consequence of a neural undersampling of the visual field.[8] As with the mechanisms underlying stereo-deficits, this suggests a potential link with acuity loss. Accordingly, some relation has been reported between acuity and crowding in adults,[7,8,32] although others report that its magnitude exceeds the scale of acuity losses.[33,34] This relationship is yet to be examined in children, and the relationship between crowding and other amblyopic deficits is unknown.

These four deficits (acuity, contrast detection, crowding, and stereo-vision) are not the only functional consequences of amblyopia, but the similarities in their proposed mechanisms led us to consider how they might covary across individuals. In particular, if each were indeed linked with alterations in either the balance of monocular cells responding to each eye, or the proportion of binocular cells, then we may expect them to covary in their magnitude. The aim of the present study was thus to examine the relationship between acuity, contrast detection, crowding, and stereoacuity in children with and without strabismus and/or anisometropia. We further sought to equate stimulus and task demands across these tasks and consequently developed a novel paradigm, tailored to children, that involves simple characters in an easy-to-understand and engaging context.

Methods

Participants

We examined 72 children, between 4 and 9 years of age (54–107 months; mean = 80.7), divided into four groups: those with strabismus ($n = 18$), anisometropia ($n = 16$), or mixed strabismus/anisometropia ($n = 19$), and controls ($n = 19$). All were tested at the Children's Eye Centre in Moorfields Eye Hospital (London, UK).

Children in the *control* group were selected to have acuity better than 0.10 logMAR (logarithm of the minimum angle of resolution), in the absence of any preexisting visual deficits. Inclusion in the *strabismic* group was on the basis of manifest heterotropia and an interocular refractive difference less than 1 diopter in any meridian. Conversely, the *anisometropic* group had an interocular refractive difference of 1 diopter or more in any meridian, without significant heterotropia. Children with both a manifest heterotropia and interocular refractive differences greater than 1 diopter were included in the *mixed* group. There was no requirement for the presence of amblyopia in the three "clinical" groups because we sought to assess a range of visual abilities. We also sought a mix of stereo-vision abilities for each group, and thus the observed frequencies of stereo-blindness for these groups should not be taken as indicative of the overall population. Children with paralytic or restrictive strabismus, as well as those with additional visual deficits (e.g., cataracts or macular dystrophies) or neurologic deficits (e.g., dyslexia or autism), were excluded. The orthoptic details for all children are presented in Supplemental Tables S1 to S4 (see Supplementary Material and Supplementary Tables S1–S4, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.12-10313/-/DCSupplemental).

Pretesting

All children underwent a full orthoptic assessment, including logMAR acuity (with Thompson v2000 software; Thompson Software Solutions, Herts, UK), TNO stereoacuity testing, ocular motility, binocular function (prism fusion range), and cover tests for near and distance fixation. The angle of any heterotropia or heterophoria was measured by prism cover test. Where required, cycloplegic refraction was performed within the 6 months prior to involvement in our study (but never immediately prior to our experiments).

Apparatus

Experiments were run using a commercial software program (MATLAB; The MathWorks, Ltd., Cambridge, UK) on a Dell PC (Dell, Round Rock, TX) running PsychToolbox.[35,36] Stimuli were presented on a liquid crystal display monitor (SyncMaster 2233RZ LCD monitor; Samsung Electronics, Seoul, South Korea), with 1680 × 1050 pixel resolution and 120-Hz refresh rate. The monitor was calibrated using a spectrophotometer (Konica Minolta Sensing Americas, Ramsey, NJ) and linearized in software, giving a maximum luminance of 92.8 cd/m^2. Because these monitors allow only 6-bit contrast resolution, bit-stealing[37] was used to obtain 442 brightness levels (approximately 8.8 bits).

Children wore stereo-shutter glasses (nVidia Corp., Santa Clara, CA), which alternated at 120 Hz (delivering images to each eye at 60 Hz) and were custom-mounted in a children's ski-mask frame to give a comfortable fit over spectacles (Fig. 1A). Children viewed the stimuli from 3 m, wearing their full correction (as required), and made verbal responses that were recorded by the experimenter using the computer keyboard. To aid in the psychophysical tasks (e.g., for acuity: "which color ghost was Vac-Man facing?"), large reference pictures of the ghosts were attached to the monitor edges. This allowed children to report either the location of a ghost or, if preferred, its color.

Stimuli and Procedures

Overall, our tasks involved five video-game characters. The centrally located target stimulus, known as Visual Acuity Man (*Vac-Man*; see Fig. 1A), was a circle with a horizontal bar missing from one half of the center, similar to a "filled-in" Landolt C. The "mouth" width was equal to one-fifth the stimulus diameter, as in the Sloan alphabet. Four ghost characters were also present, acting either as color aids in the identification of Vac-Man's orientation ("which color ghost was Vac-Man facing?") or achromatic flanker stimuli in the crowding task. The width of each ghost's "leg" gaps was also one-fifth the stimulus diameter.

Each child completed four tasks (acuity, contrast, crowding, and stereoacuity) twice. The first three tasks involved monocular presentation (once to each eye), whereas the stereoacuity task was performed once with random-pixel arrays and once with additional monocular contours. Children who did not complete all eight tasks were excluded and are not included in the tallies above (three were excluded in this way).

All children began with the *acuity task*. Here, Vac-Man was presented in the center of the display and rendered in black at 50% Weber contrast against a midgray background. Four ghosts moved slowly along the monitor boundaries (approximately 2.2–4.4° from the center), each with distinct colors (red to the right, green above, blue

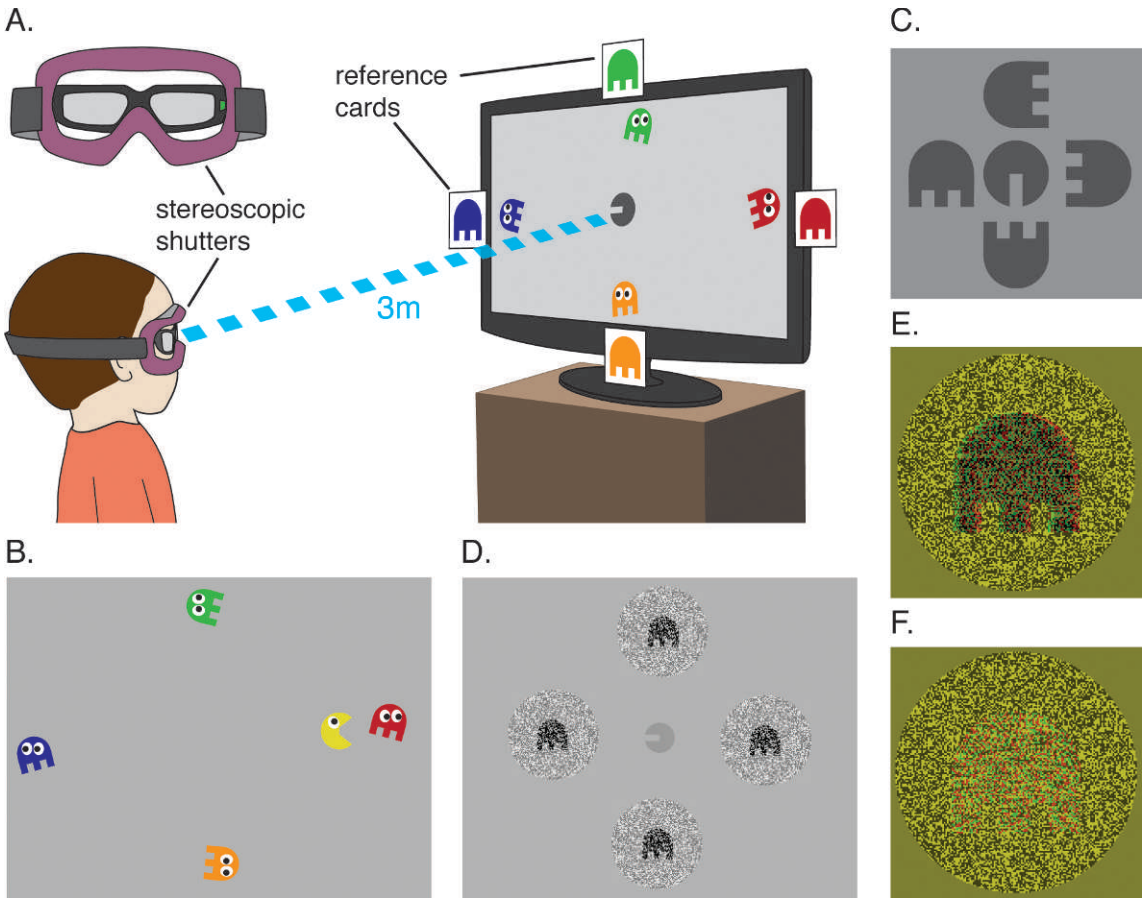


Figure 1. Apparatus and stimuli. (**A**) Stimuli were viewed through stereoscopic shutter glasses mounted in a children's ski mask (*inset*) and presented on a 3D-compatible monitor viewed at 3 m. An example trial of the acuity/contrast task is depicted. Children reported the color of the ghost that Vac-Man was facing. Colored cards were present on the monitor edges for reference. (**B**) An example frame from the "reward animation" (see Supplementary Material and Supplementary Movies S1–S4, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.12-10313/-/DCSupplemental). (**C**) Depiction of the crowding stimuli. Ghosts were rendered achromatically and presented at random orientations at a fixed separation from Vac-Man. The task was as in (**A**), with the reference cards assisting color memory. (**D**) An example frame from the stereo task with monocular "shadow" contours, representing input to one of the eyes. Children were required to help Vac-Man find the ghost that "popped out" in depth. (**E**) An example ghost from the stereo task with a monocular "shadow," viewable with *red–green* anaglyphs (*red* filter to the right eye; actual stimuli were achromatic). (**F**) An example ghost from the random-check stereo task, with a disparity offset defined purely by the random-check elements.

left, and orange below; see Fig. 1A). This combination of target-flanker separation, flanker-color, and flanker-motion minimized the chance of any crowding between target and ghosts.[9,22,38,39]

Children were required to indicate which of the four ghosts Vac-Man was facing (four-alternative forced choice), because this was the ghost "he wished to eat." These responses (based on the color of clearly individuated characters) were designed to avoid both left–right response confusions[40] and source confusions[41,42] as to which element in the multielement display was the target. Normal color-naming abilities were checked using our stimuli, prior to running the experiment. Feedback was given after each trial via brief animations (Vac-Man smiled or frowned), with every third correct response giving longer animations in which a color-rendered Vac-Man "ate" the correct ghost (Fig. 1B; see also Supplementary Material and Supplementary Movie S1, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.12-10313/-/DCSupplemental).

Acuity thresholds were assessed by varying the size of Vac-Man (and thus the visibility of the "mouth") using a QUEST algorithm staircase procedure[43] that converged on 62.5% correct performance. Incorrect responses resulted in an increase in target size and correct responses in a decrease. Three modifications were made to the QUEST routine to increase its suitability for testing children. First, children were given three practice trials with a target mouth-size of 4 minutes of arc to begin each task (although this increased if these trials were failed). Second, to minimize the frustration arising from adaptive procedures placing many trials near threshold, we presented easier catch trials on every fifth trial with stimuli scaled to double the current threshold estimate. Third, to reduce the time required to estimate thresholds, we added an exit criterion: if the SD of the estimated threshold for the past eight trials was less than 0.03 log units (a value deemed reliable during pilot tests), the experimenter was given the option of exiting the task. Otherwise, the QUEST procedure terminated after 32 trials (plus practice). The average number of trials required to reach threshold on the acuity task was 14.7, not including practice.

Children completed the acuity task once for each eye, with monocular presentation controlled using the stereo-shutter glasses. For children in the control group, their left or right eye was randomly selected to begin; those in clinical groups always began with their fellow eye to ensure comprehension. Although acuity was always completed first, subsequent tests were administered in random order. For all tasks following the acuity measurement, Vac-Man and the ghost-flankers were presented at sizes equal to 2.5-fold the size-threshold (separately for each eye) to ensure their visibility.

For the *contrast-detection* task, Vac-Man was again presented in the center of the display, surrounded by four distinctly colored ghosts, and children indicated the ghost that Vac-Man was facing (see Supplementary Material and Supplementary Movie S2, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.12-10313/-/DCSupplemental). QUEST

procedures were similar to the acuity task, albeit modulating Weber contrast instead of size, beginning at 40% Weber contrast.

For the *crowding* task, Vac-Man and ghost elements were achromatic and dark against the midgray background at 50% Weber contrast (see Fig. 1C). The ghost-flankers were rendered without their eyes and presented at random orientations. The target and flanker stimuli were thus sufficiently similar to induce crowding[9,38]; indeed, these stimuli produced strong crowding in the periphery of unaffected adults during pilot testing. Ghost-flankers were positioned in each cardinal direction at a given separation (determined by QUEST) from the centrally located Vac-Man (see Supplementary Material and Supplementary Movie S3, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.12-10313/-/DCSupplemental). As before, children indicated which ghost Vac-Man was facing, using the color of the ghosts that was also available on reference cards at the monitor edges. As above, these distinct character "identities" were intended to avoid poor performance arising from source confusions,[41,42] rather than an integrative crowding process. QUEST operated as before, but now altered the center-to-center separation between Vac-Man and ghost-flankers to converge on the minimum/critical spacing leading to accurate identification. Because crowding is largely insensitive to stimulus size,[26,44,45] critical spacing is a more reliable measure of crowding than varying size at a fixed separation. Target-flanker separation began at six times the Vac-Man radius. If identification was successful until target and flanker elements were abutting, QUEST continued for three trials and exited if all were correct, recording a zero value.

Note that these zero values reflect ceiling performance only for our element sizes (scaled to 2.5× the acuity threshold for each child). It is possible that a reduction in element size would reveal a small degree of crowding in these cases by allowing smaller interelement separations. However, because performance under such conditions is limited largely by contrast masking,[46] which operates at or near the resolution limit, we did not attempt further testing with the very small stimuli required to yield reliable estimates of performance. We also deemed the demands of this task (e.g., very high levels of fixation stability) to be beyond the children being tested. Instead, we interpret these zero values as an indication of the functional absence of foveal crowding.

Finally, we measured *stereoacuity* both with and without the presence of a monocular "shadow" contour. Ghost stimuli were rendered as random-check stereograms within a circular region. The target was now one of the four ghosts, presented with binocular disparity, whereas the others lay in the zero-disparity plane. An example screen frame is depicted in Figure 1D (also see Supplementary Material and Supplementary Movie S4, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.12-10313/-/DCSupplemental). Ghosts (without eyes) were defined by a region of checks that were correlated in both left and right eyes, with some displacement to give binocular disparity (as rendered for red/green anaglyphs in Fig. 1E). Surrounding checks were the same in each eye and random checks were used to fill the region left by interocular displacements. The random checks were light or dark squares with 50% Weber contrast.

In the "contour-based" stereograms, ghosts were additionally cued by a 25% contrast decrement on top of the existing pixel modulations (as in Fig. 1E), which followed the interocular correlation of the constituent pixels. Prior studies have shown that contour information may provide improvements to stereo-vision over purely random-pixel–based stimuli,[47–50] although it is not clear whether this reflects the use of additional monocular "cues" to the task.[51,52] In our task, this monocular "shadow" was present in all four ghosts in the display. However, to avoid this "monocular contour" being a cue to the presence of disparity on its own, distractor ghost images were selected randomly from either the right- or left-eye displacements of the target. Thus, horizontal offsets of the ghosts within the circular random-check array could not be used as a cue to the presence of the target ghost (as in Fig. 1D). In the "random-check only" stereograms, monocular shadows were absent and ghosts were defined solely by interocular pixel correlations (depicted in Fig. 1F). Thus, with sufficient stereopsis, the target ghost would be visible in one of the patches and no ghosts would be visible in the distractor patches.

In each version of the stereo task, children were instructed to "help Vac-Man find the ghost that pops off the screen." During trials, Vac-Man was centrally located at zero disparity and randomly faced each of the ghosts for brief periods. As above, practice trials were given to begin, with feedback indicating the target ghost to ensure task comprehension. Initially, each of the random checks corresponded to one monitor pixel (19 arcseconds), but if children failed to identify the target these checks were scaled with square dimensions of 2, 4, 8, or 16 monitor pixels. Failure at one scale meant an increment to the next (which typically required four to five trials), until either a satisfactory check size was found or it was clear that the child had no demonstrable stereo-vision. Once the correct scale had been selected, QUEST began with a disparity of 160 arcseconds and ran as before, varying the binocular disparity of the target ghost. Children completed the monocular-contours task first before the random-checks only version.

All procedures were approved by local and National Health Service ethics boards and conformed to the Declaration of Helsinki.

Results

Thresholds for the Spatial Tasks

We first consider performance in the three spatial tasks (acuity, contrast-detection, and crowding). The distributions of acuity and contrast thresholds were both significantly skewed and were consequently log-transformed. Crowding values were not transformed due to the high number of zero values. Thresholds for the acuity, contrast-detection, and crowding tasks are presented in Figure 2. Data from the clinical groups are attributed to either the "amblyopic" or "fellow eye"; control group data are kept as left/right eyes. Colored bars depict group means, whereas gray points show individual data.

Acuity thresholds (Fig. 2A) for the left and right eyes of the control group averaged 0.03 and 0.01 log minutes of arc, respectively, both equivalent to a Snellen acuity of 6/6. There was no significant difference between these values (paired differences *t*-test: $t[18] = 0.66$, $P = 0.5$). Reduced acuity levels were evident in the amblyopic eyes of all three clinical groups, with averages of 0.32, 0.23, and 0.32 log minutes for the strabismic, anisometropic, and mixed groups, respectively (equivalent to Snellen acuities of 6/12 and 6/10), compared with average fellow-eye acuities of 0.10, –0.01, and 0.02 log minutes. These interocular differences were significant for each group (paired differences for strabismic: $t[17] = 4.6$, $P < 0.001$; anisometropic: $t[15] = 5.9$, $P < 0.001$; and mixed: $t[18] = 9.4$, $P < 0.001$). This is solely due to elevations in the amblyopic eye—acuity in the fellow eyes of these groups was not significantly different from that of controls (pooled unpaired *t*-test between all fellow eyes and both control eyes: $t[89] = -0.6$, $P = 0.56$). These acuity values also correlate highly with logMAR values obtained during orthoptic examination (see Supplementary Material and Supplementary Fig. S1A, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.12-10313/-/DCSupplemental) and show good agreement in a Bland–Altman analysis (see Supplementary Material and Supplementary Fig. S1B, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.12-10313/-/DCSupplemental).

Contrast-detection thresholds (Fig. 2B) averaged 15.9% and 14.8% Weber contrast for the left and right eyes of controls, a nonsignificant difference ($t[18] = 0.5$, $P = 0.5$). Group averages in the clinical groups showed slight but nonsignificant elevations in the amblyopic eye: strabismic children averaged 19.2% and 17.7% for the amblyopic and fellow eyes ($t[17] = 0.5$, $P = 0.49$), anisometropes were 17.3% and 14.0% ($t[15] =$

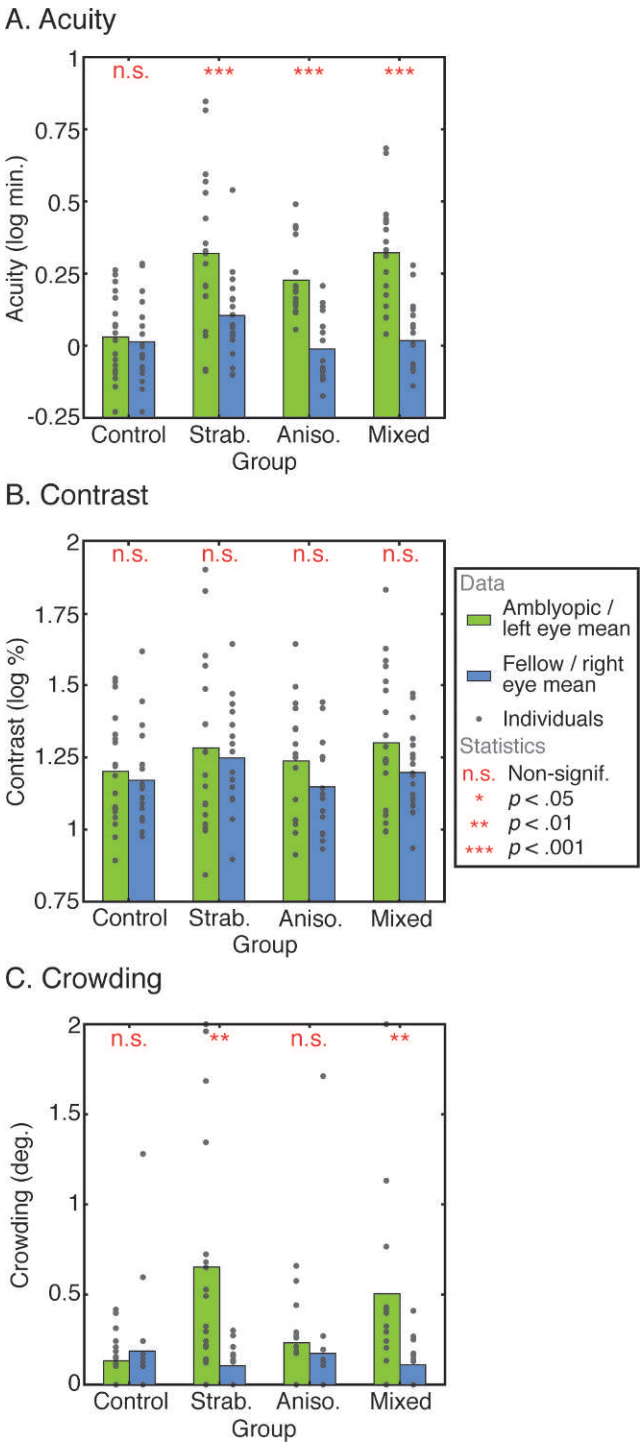


Figure 2. Thresholds for the three spatial tasks. (**A**) Acuity thresholds, in log minutes of visual arc (equivalent to logMAR). Data are separated into left- and right-eye thresholds for the controls and amblyopic/fellow eye for the three clinical groups. *Colored bars* show group means; *gray points* are individual data. *Symbols* at the *top* show the significance of paired-samples *t*-tests between the eyes of group members (as in the figure *legend*). (**B**) Contrast-detection thresholds, in log units of percent Weber contrast, expressed as in (**A**). (**C**) Thresholds for the spatial extent of foveal crowding, expressed as the center-to-center separation between target and flankers in degrees of visual arc.

1.5, $P = 0.16$), and the mixed group difference approached significance with 20.0% and 15.8% ($t[18] = 1.97$, $P = 0.06$).

On average, foveal crowding was evident in both eyes of controls (Fig. 2C), with threshold center-to-center separations of 0.13° and 0.18° required for left and right eyes, respectively (a nonsignificant difference, $t[18] = -0.97$, $P = 0.34$). These are large values relative to acuity: control children had an average acuity of 1 arc minute, whereas the average crowding extent is ±9 arc minutes around this. Although the amblyopic eye of anisometropes was slightly elevated, with 0.23° compared with 0.17° in their fellow eyes, this difference was nonsignificant ($t[18] = 3.03$, $P = 0.53$). Large significant differences, however, were evident in the strabismic group, with an average separation of 0.66° required in the amblyopic eye and 0.10° in the fellow eye ($t[17] = 3.7$, $P = 0.002$), and similarly in the mixed group with 0.51° and 0.11°, respectively ($t[18] = 3.03$, $P = 0.007$).

Note that in the fellow eye of the strabismic and mixed groups, and both eyes of anisometropes, the level of crowding resembles that of the control children rather than a zero mean. Indeed, a one-way *t*-test shows a significantly nonzero mean for these eyes (pooled one-way analysis for all control and fellow-eye thresholds: $t[90] = 5.63$, $P < 0.001$). A pooled unpaired *t*-test between control eyes and fellow eyes was also nonsignificant ($t[89] = 0.63$, $P = 0.53$), demonstrating the similarity in these values between the two groups. It would appear that the fellow eyes of these children follow a similar developmental trajectory to those of controls. There was, however, no influence of age across the approximately 50-month range of our sample (see Supplementary Material and Supplementary Fig. S3, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.12-10313/-/DCSupplemental).

Correlations between the Spatial Tasks

To examine the relationship between the interocular differences (IODs) shown above, we first converted monocular thresholds to IOD values by subtracting right-eye from left-eye thresholds in the control group and fellow-eye from amblyopic-eye thresholds in the clinical groups. Intertask comparisons of these IODs are plotted in Figure 3. Negative IOD values indicate better performance in the amblyopic eye (or right eye for controls) than the fellow eye (or left eye), and vice versa for positive values. The clustering of data points in Figure 3A shows that there is no clear relationship between the IODs in acuity and contrast detection for any of the four groups (Control: $r[17] = -0.02$, $P = 0.94$; strabismus: $r[16] = 0.15$, $P = 0.55$; anisometropia: $r[14] = 0.26$, $P = 0.33$; mixed: $r[17] = 0.21$, $P = 0.38$). Similarly, Figure 3B demonstrates the lack of significant correlations between IODs for contrast detection and crowding (Control: $r[17] = -0.07$, $P = 0.78$; strabismus: $r[16] = 0.26$, $P = 0.29$; anisometropia: $r[14] = 0.15$, $P = 0.59$; mixed: $r[17] = 0.22$, $P = 0.37$).

Figure 3C shows the much stronger relationship between acuity and crowding. Positive correlations are evident in all groups, with larger interocular differences in acuity associated with larger interocular differences in crowding. This was significant for the control ($r[17] = 0.54$, $P = 0.02$), strabismic ($r[16] = 0.59$, $P = 0.01$), and anisometropic ($r[14] = 0.52$, $P = 0.04$) groups, and approached significance in the mixed group ($r[17] = 0.43$, $P = 0.07$). The relationship is such that for strabismic children, for instance, a 1-minute increase in the IOD for acuity (without log conversion) corresponds to a 0.24° increase in the IOD for crowding. So, although the control and anisometropic groups do not show significant IODs in crowding at the group level (Fig. 2), the small IODs that do occur are nonetheless well predicted by their interocular acuity differences. The larger crowding IODs of the strabismic

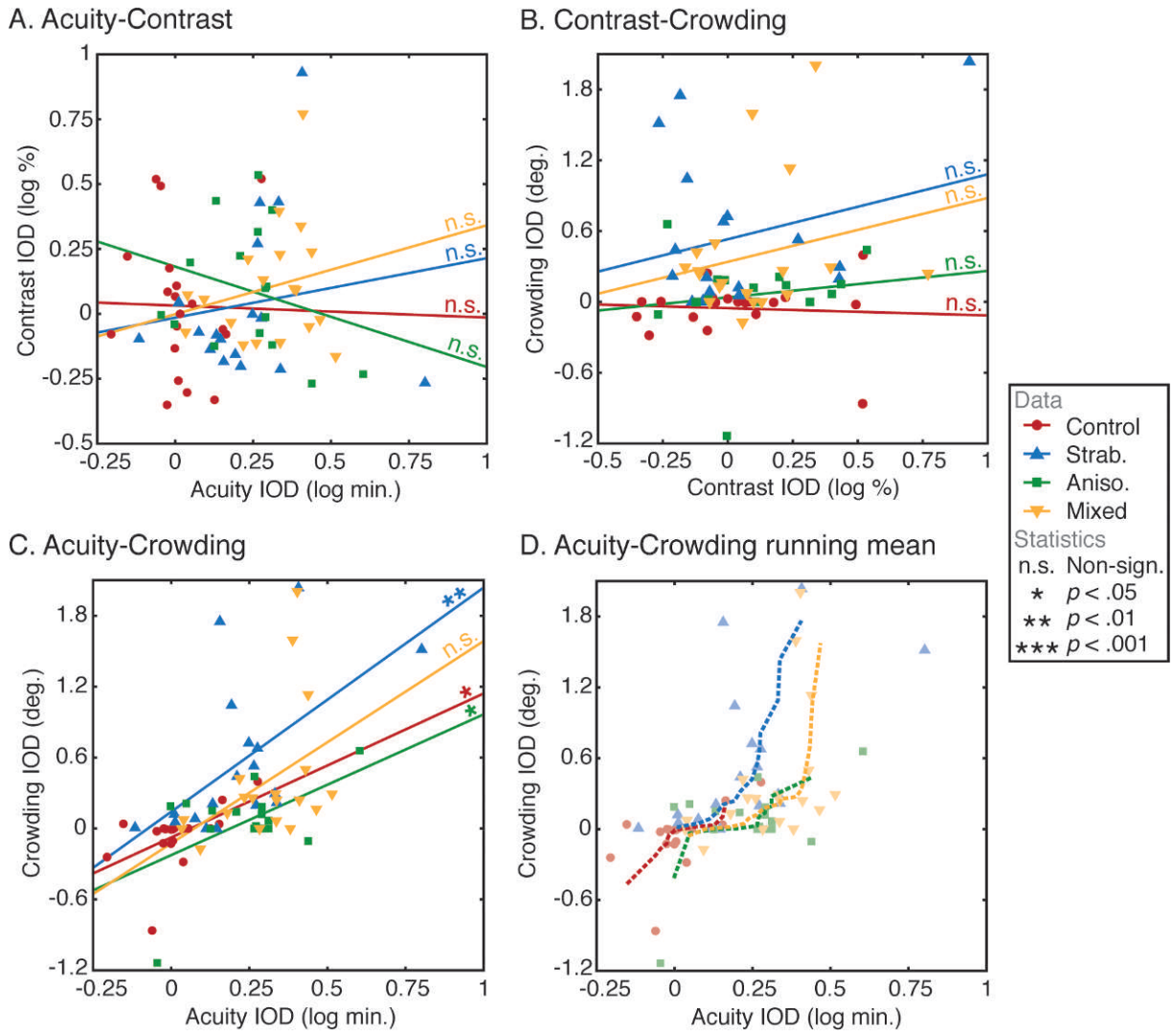


Figure 3. Correlations between interocular differences on each of the three spatial tasks. (**A**) Correlations between interocular differences in acuity (log minutes difference) and contrast detection (log percent difference). Individual *points* show each observer and lines for each group display the best-fitting line to the data, with significance indicated on the *rightward* end. (**B**) Correlations between IODs in contrast detection and crowding, plotted as in (**A**). (**C**) Correlations between IODs for acuity and crowding (degrees difference), plotted as in (**A**). (**D**) Acuity and crowding IODs are replotted, with running means (±1 child) for the crowding IODs shown as a function of the acuity IOD. Data *points* (as in **C**) are dimmed for clarity.

and mixed groups are similarly well predicted by interocular acuity differences, although the nonsignificant correlation for the mixed group appears to arise from those with the largest acuity IODs, whose crowding IODs were disproportionately large. Accordingly, Figure 3D plots the running mean for crowding IOD values for a given child with their nearest neighbors (±1 child) sorted by the size of acuity IODs. For the strabismic and mixed groups there is a sharp rise in crowding IODs once acuity differences rise above 0.25 to 0.35 log minutes. In contrast, control and anisometropic values show a linear relationship throughout their ranges.

Note that the lack of correlation between contrast sensitivity and either acuity or crowding, reported above, is important since it demonstrates that children performing well on one test did not necessarily mean they would perform well on another. This rules out any explanation of the substantial intertask correlations between acuity and crowding based on general cognitive factors such as attention.

Stereoacuity Thresholds

We consider performance on the stereo tasks separately here because of the significant number of children who failed to achieve measurable stereoacuity values. Children were considered to have "passed" these tasks if their thresholds were below 1000 arcseconds, a value well above the range of stereoabilities expected for children of this age.[10,14,51,53] For each group, the percentage of children who passed each stereo task is plotted in Figure 4A. In each case, more children achieved measurable stereoacuity thresholds in the monocular-contours–based task than in the random-checks–only version (control: 18/19 vs. 16/19; strabismic: 5/18 vs. 2/18; anisometropic: 13/16 vs. 10/16; and mixed: 5/19 vs. 2/19).

To examine the precise stereoacuity thresholds obtained, we next consider only the children with measurable stereoacuity on *both* stereo tasks (i.e., the latter proportions for each group above). Because this reduced the number of children in the clinical groups significantly, we compare our control children with a pooled "clinical" group consisting of children

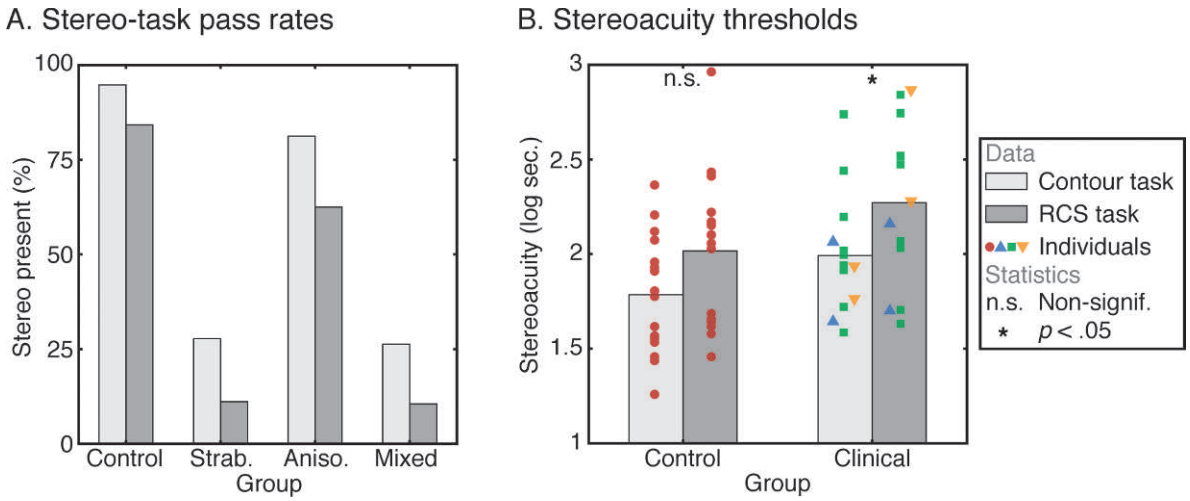


Figure 4. Performance on the two stereo tasks. (**A**) The percentage of children from each group who achieved measurable stereoacuity thresholds (less than 1000 arcseconds) on either of the contour-based (*light gray bars*) or the random-check stereo task (RCS task; *dark gray bars*). (**B**) Thresholds on the two stereo tasks, plotted in log seconds of arc. Data are taken only from children with measurable stereo on *both* tasks, grouped into either "control" or "clinical" categories. The latter includes children from the strabismic (*blue triangles*), anisometropic (*green squares*), and mixed (*yellow inverted triangles*) groups. *Gray bars* show the mean thresholds for each group on each task. The results of paired differences *t*-tests are shown for each group at the *top* of the figure, with significance levels shown in the figure *legend*.

from each of the three categories. As with the acuity and contrast-detection thresholds, stereoacuity thresholds were significantly skewed and were thus log-transformed, as presented in Figure 4B. Control children achieved average stereoacuities of 61 and 104 arcseconds on the contour and random-check tasks, respectively, whereas children in the clinical group achieved thresholds of 98 and 187 arcseconds on the two tasks. A two-way ANOVA gave a significant main effect of group ($F[1,56] = 5.92$, $P = 0.018$), reflecting the higher thresholds in the clinical group, and a significant main effect of stereo task ($F[1,56] = 7.25$, $P = 0.009$), with a nonsignificant interaction ($F[1,56] = 0.06$, $P = 0.80$). That is, both groups performed better on the contours task than the random-check stereo task, although clinical children were worse overall. These values correlate significantly with TNO stereoacuity thresholds obtained during orthoptic examination and show modest agreement in Bland–Altman analyses (see Supplementary Material and Supplementary Fig. S2, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.12-10313/-/DCSupplemental).

The Presence of Stereo-Vision

To examine the relationship between stereoacuity and the three spatial tasks, we sought to include those without measurable stereo-thresholds in the analysis, but without assigning these children arbitrary stereo-thresholds (which could potentially bias correlation analyses). We therefore divided children into two groups: "stereo," for those with measurable thresholds (below 1000 arcseconds) on *either* of our Vac-Man stereo tasks, and "nonstereo" for those who failed both tasks. Children were grouped in this way regardless of their clinical condition. Note that we did not incorporate data from the control group simply to avoid them dominating the "stereo" category. This gave a population of 23 stereo and 30 nonstereo observers, whose thresholds are plotted in Figure 5. Individual points show observers grouped by their clinical condition, whereas gray bars show overall means.

Figure 5A plots the IODs in acuity, which are smaller in the stereo group (an average 0.2 log minute difference) than the nonstereo group (0.3 log minute), a significant difference (unpaired *t*-test: $t[51] = -2.07$, $P = 0.04$). Inspection of the individual points reveals this pattern in each of the three clinical groups. In contrast, interocular differences for contrast detection (Fig. 5B) were similar for stereo and nonstereo individuals, and did not differ significantly ($t[51] = -0.87$, $P = 0.39$), although there was a tendency for nonstereo strabismic children to have improved contrast detection in their amblyopic eye.

Crowding IODs show a large difference between the groups (Fig. 5C), with small IODs for stereo children (0.05° on average) and large IODs for nonstereo children (0.58° on average), a significant difference ($t[51] = -3.78$, $P < 0.001$). Although this could be due to the predominance of strabismic and mixed amblyopes in the nonstereo group, the same pattern is clearly evident within each of the three clinical categories. Interestingly, the sole control subject who failed both our stereo tasks also showed elevated crowding with an IOD of –0.86°, much larger than the average control IOD of –0.01°. A similar pattern is evident within the "stereo" category of children when precise stereoacuity thresholds are plotted against these IOD values (see Supplementary Material and Supplementary Fig. S4, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.12-10313/-/DCSupplemental), although the reduced range of these values and lower subject numbers considerably reduces the clarity of these effects.

Discussion

We report four key findings. The first two document the threshold abilities of children in the control and clinical groups, and replicate previous results, while the second two present novel insights into the correlation between these thresholds.

Visual Abilities in Childhood and Amblyopia

First, our *Vac-Man* paradigm successfully detected the interocular acuity differences of strabismic, anisometropic, and mixed amblyopia, as well as interocular differences in foveal crowding for the strabismic and mixed groups. This pattern of crowding deficits is consistent with previous observations in childhood[28,30,31] and adult[7–9] amblyopia, and rises above the already significant levels of crowding that we

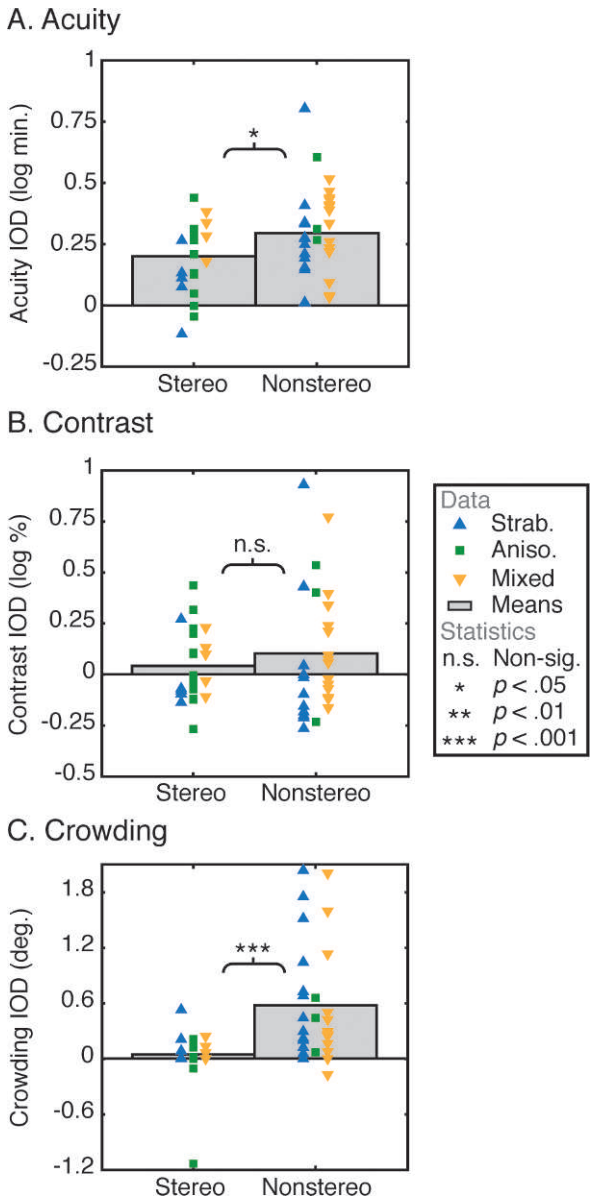


Figure 5. Interocular differences on the three spatial tasks for stereo and nonstereo observers in the three clinical groups. (**A**) IODs in acuity, plotted in log minutes difference. *Points* show individual observers (colored according to their clinical classification), whereas group averages are shown as *gray bars*. The result of a two-way *t*-test between stereo and nonstereo observers is indicated by the central bracket, with significance levels shown in the figure *legend*. (**B**) IODs in contrast-detection thresholds, in log percent difference units, plotted as in (**A**). (**C**) IODs in crowding, in degrees of difference, plotted as in (**A**).

observe in both eyes of all children (as observed previously[27–29]). Here we demonstrate crowding using a task with clearly individuated elements (that are nonetheless sufficiently similar to induce crowding), thus avoiding the possibility of poor performance due to source confusions[41,42] or miscomprehension of the task.

Our second key finding is that children had higher pass rates and lower stereoacuity thresholds in a stereo task with monocular contours than a random-checks–only version of the same task. This is consistent with prior observations,[47–50] although others have argued that these improvements reflect the use of monocular cues.[51,52] That is, the contours could provide either spatial displacement cues (e.g., if the element offset in depth was clearly displaced leftward in one eye) or motion parallax cues during head movement. In our task, the use of horizontal displacement cues was eliminated by randomly sampling the zero-disparity distractors from either of the left- or right-eye stereo-images. Motion parallax cues should also have been minimal given the 3-m viewing distance.[54] We suggest instead that the contours may serve as a vergence cue to guide more efficient fusion of the two images.[47,50] It is also possible that the larger contours aid stereopsis because of their additional low spatial frequency content, whereas random-check stimuli rely exclusively on high spatial frequencies. In anisometropic adults, stereoacuity indeed shows greater impairments in the high spatial frequency range.[55] This finding not only suggests that amblyopic stereoacuities may be systematically underestimated, but also that stereo-training programs may have greater success if the initial stages use contour-rich stimuli.

Unlike previous studies, we did not observe any differences in contrast-detection sensitivity between our groups. The lack of effect on contrast detection in children differs from adult amblyopia, where the amblyopic eye of anisometropes and mixed strabismic–anisometropes typically shows impaired contrast detection, whereas that of strabismics often improves.[2] We report a trend toward this pattern in our data, with the majority of our strabismic group showing better contrast detection in the amblyopic eye (67%) compared with more even distributions in the anisometropic (50%) and mixed (42%) groups. The changes in the visual system that produce this pattern of improvements and deficits may simply take longer to develop than the effects on acuity and crowding observed herein. Accordingly, although contrast deficits are evident in adults with broadband stimuli[4,5] (albeit at different magnitudes depending on the spatiotemporal frequency and color content[6]), deficits have been observed in children only with high spatial-frequency stimuli.[3] The use of narrowband stimuli may more closely reveal the development of these effects.

Crowding and Its Relation to Acuity

Our third key finding is that interocular differences in acuity and crowding correlate for children in all groups (including control and anisometropic children, albeit trending in the mixed group). This suggests that childhood deficits in acuity and crowding are linked, regardless of their magnitude. Importantly, this correlation is not simply a general covariation in psychophysical performance, as contrast-detection thresholds did not correlate with either acuity or crowding.

The correlation between acuity and crowding in strabismic children is consistent with that shown previously in strabismic adults.[7,8,32] Others, however, have found that the scale of foveal crowding in strabismic adults can greatly exceed that of their acuity losses.[33,34] In fact, our data show both features: although acuity and crowding IODs correlate for all children, crowding IODs in the strabismic and mixed groups increase sharply once acuity IODs exceed some critical value (Fig. 3D). This additional increase in crowding more closely resembles the excessive scale of crowding observed in cases of adult strabismus[33,34] and was greatest in the mixed group, resulting in the nonsignificant trend between acuity and crowding for these children. Further, the few anisometropes in our sample with large acuity IODs do not exhibit this sharp elevation in crowding, suggesting that this additive crowding may require the presence of strabismus. Were the strength of this additive element to increase over time, it could produce the dissociation between acuity and crowding observed in some

adult amblyopes.[33,34] This differs from the "normal" adult fovea, where acuity and crowding continue to be correlated,[7] consistent with our observations in unaffected children. Although the magnitude of foveal crowding decreases significantly in adulthood,[29] the correlation with acuity thus appears to remain. It could be the case that acuity and crowding are initially linked, before diverging in strabismic amblyopia (if unsuccessfully treated) as the deficit builds over time.

A potential complication with the observed link between acuity and crowding arises from our variation in stimulus sizes, which were scaled relative to acuity thresholds. Although crowding is largely size invariant in both the strabismic fovea[34] and the normal periphery,[44,56] in the normal adult fovea it varies with size.[46] Our observed covariations between acuity and the extent of foveal crowding could therefore simply result from the larger stimulus sizes used to measure crowding. Although the size tuning of foveal crowding in children is not known, dividing the spatial extent of crowding by acuity thresholds nonetheless reveals a persistent degree of crowding in all groups (see Supplementary Material and Supplementary Fig. S5A, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.12-10313/-/DCSupplemental). The increase in crowding IODs for nonstereo children over those with intact stereo-vision is similarly resistant to this correction (see Supplementary Material and Supplementary Fig. S5B, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.12-10313/-/DCSupplemental). Since stimulus size and acuity were directly proportional in our study, it is not possible to determine whether size-corrected crowding values continue to correlate with acuity, but given the above we suggest that these correlations are unlikely to solely reflect variations in stimulus size.

Stereo-Vision and Its Relation to Acuity and Crowding

Our fourth key finding is that interocular differences in both acuity and crowding were significantly larger for children without measurable stereoacuity than for those with measurable stereoacuity. Correlations between these abilities were also evident in children with measurable stereoacuity thresholds (see Supplementary Material and Supplementary Fig. S4, http://www.iovs.org/lookup/suppl/doi:10.1167/iovs.12-10313/-/DCSupplemental). Together, these results suggest a link between acuity, crowding, and stereo-vision both in normal development and in cases of strabismus and anisometropia. This link between large IODs in acuity and the loss of stereo-vision has been observed previously in adults,[2,10,12] but to our knowledge, the present study is the first demonstration of a link between stereo-vision and crowding.

The correlation between these three deficits suggests that they may share a common cause. Accordingly, both interocular acuity differences and the loss of stereoacuity have previously been attributed to a loss of binocular cells and a shift in the selectivity of monocular cells away from the amblyopic eye.[1,2,16–18] The loss of binocular cells would obviously impair stereo-vision, while this in conjunction with the shift in monocular cell selectivities could also impair acuity. That is, with monocular viewing, the amblyopic eye of a nonstereo observer would have access only to its own monocular cells, whereas the fellow eye would also have access to the formerly binocular cells. In addition, this redirection of cell selectivities could also lead to elevations in crowding; if crowding were the result of a neural undersampling of the visual field,[8] then the reduced sampling of the amblyopic eye may prompt an increase in this pooling process. In the normally developing visual system, the maturation of these cell selectivities may give similar linkages between these abilities.

Alternately, both strabismic and anisometropic amblyopia have been linked with an increase in positional noise,[57–59] which could certainly reduce the visibility of fine detail and hinder the detection of the interocular correlations in stereo-vision. Positional uncertainty has also been proposed to account for crowding,[60,61] although weighted averaging provides a closer approximation of the systematic errors that arise in crowded conditions.[23,24] It is possible that positional noise may contribute to these processes in the case of amblyopia, however.[62] Age-related improvements are also seen for Vernier acuity in normally developing children,[63,64] suggesting that the maturation of positional encoding processes could similarly underlie the linked impairments in acuity, crowding, and stereo-vision. Of course, these two mechanisms (alterations in cell selectivity and increased positional noise) are not mutually exclusive and some combination of the two is possible.

The Neural Loci of These Abilities

If acuity, crowding, and stereo-vision do share a common basis, as we suggest above, where could the neural changes that give rise to these covariations occur? An obvious candidate is V1, which changes substantially under amblyopic conditions.[1,18] As above, shifts in the ocular dominance of V1 cells could account for both the acuity losses and the disruption of stereo-vision observed herein. However, it has been noted that the scale of psychophysical deficits in amblyopia does not match the scale of physiologic changes in V1.[65] Changes beyond V1 are therefore likely, although precisely where these changes occur is unclear.

The neural basis of crowding is even more elusive. It has long been thought to occur at least at the level of V1 binocular cells,[66] although many psychophysical effects have suggested that its primary operation may be at higher levels.[26,67–70] Recent work suggests that crowding may in fact operate at multiple levels in the visual hierarchy,[71] increasingly modulating activity throughout cortical regions V1 to V4.[72] The multilevel nature of crowding could mean that some of the elevated crowding we observe results from changes in early visual cortex (e.g., V1). These early-stage deficits could produce the covariation with acuity and stereo-vision, although of course this could also arise if these alterations in V1 selectivity were inherited by higher-level regions. In contrast, the apparent dissociation of crowding from acuity with large IODs (in the presence of strabismus) could reflect distinct processes beyond V1. Similarly, although the developmental forms of crowding we observe in our control group could reflect the maturation of V1 function, the late maturation of connectivity and function in higher cortical regions[73] suggests that later-stage processes could be what drive these elevations. The neural changes that give rise to elevated crowding may thus be widespread throughout the visual system. Regardless of their origins, however, our results suggest that these changes have consistent effects on acuity, crowding, and stereo-vision.

Acknowledgments

The authors thank Kate Naylor, Emma Jenkinson, Marko Nardini, Eliza Burton, and Jennifer Bales for help with recruitment, and Elaine Anderson and Gary Rubin for helpful discussion.

References

1. Barrett BT, Bradley A, McGraw PV. Understanding the neural basis of amblyopia. *Neuroscientist*. 2004;10:106–117.
2. McKee SP, Levi DM, Movshon JA. The pattern of visual deficits in amblyopia. *J Vis*. 2003;3:380–405.

3. Howell ER, Mitchell DE, Keith CG. Contrast thresholds for sine gratings of children with amblyopia. *Invest Ophthalmol Vis Sci*. 1983;24:782–787.
4. Levi DM, Harwerth RS. Spatio-temporal interactions in anisometropic and strabismic amblyopia. *Invest Ophthalmol Vis Sci*. 1977;16:90–95.
5. Hess RF, Howell ER. The threshold contrast sensitivity function in strabismic amblyopia: evidence for a two type classification. *Vision Res*. 1977;17:1049–1055.
6. Davis AR, Sloper JJ, Neveu MM, Hogg CR, Morgan MJ, Holder GE. Differential changes of magnocellular and parvocellular visual function in early- and late-onset strabismic amblyopia. *Invest Ophthalmol Vis Sci*. 2006;47:4836–4841.
7. Flom MC, Weymouth FW, Kahneman D. Visual resolution and contour interaction. *J Opt Soc Am*. 1963;53:1026–1032.
8. Levi DM, Klein SA. Vernier acuity, crowding and amblyopia. *Vision Res*. 1985;25:979–991.
9. Levi DM. Crowding—an essential bottleneck for object recognition: a mini-review. *Vision Res*. 2008;48:635–654.
10. Weakley DR. The association between nonstrabismic anisometropia, amblyopia, and subnormal binocularity. *Ophthalmology*. 2001;108:163–171.
11. Goodwin RT, Romano PE. Stereoacuity degradation by experimental and real monocular and binocular amblyopia. *Invest Ophthalmol Vis Sci*. 1985;26:917–923.
12. Levi DM, McKee SP, Movshon JA. Visual deficits in anisometropia. *Vision Res*. 2011;51:48–57.
13. Cooper J, Feldman J. Random-dot-stereogram performance by strabismic, amblyopic, and ocular-pathology patients in an operant-discrimination task. *Am J Optom Physiol Opt*. 1978; 55:599–609.
14. Leske DA, Holmes JM. Maximum angle of horizontal strabismus consistent with true stereopsis. *J Am Assoc Pediatr Ophthalmol Strabismus*. 2004;8:28–34.
15. Dobson V, Miller JM, Clifford-Donaldson CE, Harvey EM. Associations between anisometropia, amblyopia, and reduced stereoacuity in a school-aged population with a high prevalence of astigmatism. *Invest Ophthalmol Vis Sci*. 2008; 49:4427–4436.
16. Wiesel TN, Hubel DH. Single-cell responses in striate cortex of kittens deprived of vision in one eye. *J Neurophysiol*. 1963;26: 1003–1017.
17. Marshman WE, Dawson E, Neveu MM, Morgan MJ, Sloper JJ. Increased binocular enhancement of contrast sensitivity and reduced stereoacuity in Duane syndrome. *Invest Ophthalmol Vis Sci*. 2001;42:2821–2825.
18. Kiorpes L, McKee SP. Neural mechanisms underlying amblyopia. *Curr Opin Neurobiol*. 1999;9:480–486.
19. Knox PJ, Simmers AJ, Gray LS, Cleary M. An exploratory study: prolonged periods of binocular stimulation can provide an effective treatment for childhood amblyopia. *Invest Ophthalmol Vis Sci*. 2012;53:817–824.
20. Hess RF, Mansouri B, Thompson B. A binocular approach to treating amblyopia: antisuppression therapy. *Optom Vis Sci*. 2010;87:697–704.
21. Ding J, Levi DM. Recovery of stereopsis through perceptual learning in human adults with abnormal binocular vision. *Proc Natl Acad Sci USA*. 2011;108:E733–E741.
22. Bouma H. Interaction effects in parafoveal letter recognition. *Nature*. 1970;226:177–178.
23. Parkes L, Lund J, Angelucci A, Solomon JA, Morgan M. Compulsory averaging of crowded orientation signals in human vision. *Nat Neurosci*. 2001;4:739–744.
24. Greenwood JA, Bex PJ, Dakin SC. Positional averaging explains crowding with letter-like stimuli. *Proc Natl Acad Sci USA*. 2009;106:13130–13135.
25. Greenwood JA, Bex PJ, Dakin SC. Crowding changes appearance. *Curr Biol*. 2010;20:496–501.
26. Pelli DG, Palomares M, Majaj NJ. Crowding is unlike ordinary masking: distinguishing feature integration from detection. *J Vis*. 2004;4:1136–1169.
27. Atkinson J, Braddick OJ. Assessment of visual acuity in infancy and early childhood. *Acta Ophthalmol Scand Suppl*. 1983; 157(suppl):18–26.
28. Atkinson J, Anker S, Evans C, McIntyre A. The Cambridge Crowding Cards for preschool visual acuity testing. In: *Transactions of the 6th International Orthoptic Congress, Harrogate, UK*. 1987:482–486.
29. Jeon ST, Hamid J, Maurer D, Lewis TL. Developmental changes during childhood in single-letter acuity and its crowding by surrounding contours. *J Exp Child Psychol*. 2010;107:423–437.
30. Hilton AF, Stanley JC. Pitfalls in testing children's vision by the Sheridan Gardiner single optotype method. *Br J Ophthalmol*. 1972;56:135–139.
31. Simmers AJ, Gray LS, Spowart K. Screening for amblyopia: a comparison of paediatric letter tests. *Br J Ophthalmol*. 1997; 81:465–469.
32. Bonneh YS, Sagi D, Polat U. Local and non-local deficits in amblyopia: acuity and spatial interactions. *Vision Res*. 2004; 44:3099–3110.
33. Hess RF, Dakin SC, Tewfik M, Brown B. Contour interaction in amblyopia: scale selection. *Vision Res*. 2001;41:2285–2296.
34. Levi DM, Hariharan S, Klein SA. Suppressive and facilitatory spatial interactions in amblyopic vision. *Vision Res*. 2002;42: 1379–1394.
35. Brainard DH. The psychophysics toolbox. *Spat Vis*. 1997;10: 433–436.
36. Pelli DG. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis*. 1997;10:437–442.
37. Tyler CW, Chan H, Liu L, McBride B, Kontsevich LL. Bit stealing: how to get 1786 or more gray levels from an 8-bit color monitor. In: *Human Vision, Visual Processing, and Digital Display III*. Bellingham, WA: SPIE Press;1992;1666: 351–364.
38. Kooi FL, Toet A, Tripathy SP, Levi DM. The effect of similarity and duration on spatial interaction in peripheral vision. *Spat Vis*. 1994;8:255–279.
39. Bex PJ, Dakin SC. Spatial interference among moving targets. *Vision Res*. 2005;45:1385–1398.
40. Corballis MC. The left-right problem in psychology. *Can Psychol*. 1974;15:16–33.
41. Chastain G. Feature mislocalizations and misjudgments of intercharacter distance. *Psychol Res*. 1982;44:51–65.
42. Goldberg MC, Maurer D, Lewis TL. Developmental changes in attention: the effects of endogenous cueing and of distractors. *Dev Sci*. 2001;4:209–219.
43. Watson AB, Pelli DG. QUEST: a Bayesian adaptive psychometric method. *Percept Psychophys*. 1983;33:113–120.
44. Levi DM, Hariharan S, Klein SA. Suppressive and facilitatory spatial interactions in peripheral vision: peripheral crowding is neither size invariant nor simple contrast masking. *J Vis*. 2002;2:167–177.
45. Strasburger H, Harvey LO, Rentschler I. Contrast thresholds for identification of numeric characters in direct and eccentric view. *Percept Psychophys*. 1991;49:495–508.
46. Levi DM, Klein SA, Hariharan S. Suppressive and facilitatory spatial interactions in foveal vision: foveal crowding is simple contrast masking. *J Vis*. 2002;2:140–166.
47. Frisby JP, Mein J, Saye A, Stanworth A. Use of random-dot stereograms in the clinical assessment of strabismic patients. *Br J Ophthalmol*. 1975;59:545–552.

48. Richards W. Stereopsis with and without monocular contours. *Vision Res*. 1977;17:967–969.
49. Garnham L, Sloper JJ. Effect of age on adult stereoacuity as measured by different types of stereotest. *Br J Ophthalmol*. 2006;90:91–95.
50. Saye A, Frisby JP. The role of monocularly conspicuous features in facilitating stereopsis from random-dot stereograms. *Perception*. 1975;4:159–171.
51. Fawcett SL, Birch EE. Validity of the Titmus and Randot circles tasks in children with known binocular vision disorders. *J Am Assoc Pediatr Ophthalmol Strabismus*. 2003;7:333–338.
52. Simons K. A comparison of the Frisby, Random-Dot E, TNO, and Randot Circles stereotests in screening and office use. *Arch Ophthalmol*. 1981;99:446–452.
53. Simons K. Stereoacuity norms in young children. *Arch Ophthalmol*. 1981;99:439–445.
54. Ono ME, Rivest J, Ono H. Depth perception as a function of motion parallax and absolute-distance information. *J Exp Psychol Hum Percept Perform*. 1986;12:331–337.
55. Holopigian K, Blake R, Greenwald MJ. Selective losses in binocular vision in anisometropic amblyopes. *Vision Res*. 1986;26:621–630.
56. Tripathy SP, Cavanagh P. The extent of crowding in peripheral vision does not scale with target size. *Vision Res*. 2002;42:2357–2369.
57. Hess RF, Field DJ. Is the increased spatial uncertainty in the normal periphery due to spatial undersampling or uncalibrated disarray? *Vision Res*. 1993;33:2663–2670.
58. Neri P, Levi DM. Spatial resolution for feature binding is impaired in peripheral and amblyopic vision. *J Neurophysiol*. 2006;96:142–153.
59. Watt RJ, Hess RF. Spatial information and uncertainty in anisometropic amblyopia. *Vision Res*. 1987;27:661–674.
60. Wolford G. Perturbation model for letter identification. *Psychol Rev*. 1975;82:184–199.
61. Strasburger H. Unfocussed spatial attention underlies the crowding effect in indirect form vision. *J Vis*. 2005;5:1024–1037.
62. Li RW, Levi DM. Characterizing the mechanisms of improvement for position discrimination in adult amblyopia. *J Vis*. 2004;4:476–487.
63. Carkeet A, Levi DM, Manny RE. Development of Vernier acuity in childhood. *Optom Vis Sci*. 1997;74:741–750.
64. Zanker JM, Mohn G, Weber U, Zeitler-Driess K, Fahle M. The development of Vernier acuity in human infants. *Vision Res*. 1992;32:1557–1564.
65. Kiorpes L, Kiper DC, O'Keefe LP, Cavanaugh JR, Movshon JA. Neuronal correlates of amblyopia in the visual cortex of macaque monkeys with experimental strabismus and anisometropia. *J Neurosci*. 1998;18:6411–6424.
66. Flom MC, Heath GG, Takahashi E. Contour interaction and visual resolution: contralateral effects. *Science*. 1963;142:979–980.
67. He S, Cavanagh P, Intriligator J. Attentional resolution and the locus of visual awareness. *Nature*. 1996;383:334–337.
68. Dakin SC, Greenwood JA, Carlson TA, Bex PJ. Crowding is tuned for perceived (not physical) location. *J Vis*. 2011;11:1–13.
69. Maus GW, Fischer J, Whitney D. Perceived positions determine crowding. *PLoS ONE*. 2011;6:e19796.
70. Greenwood JA, Bex PJ, Dakin SC. Crowding follows the binding of relative position and orientation. *J Vis*. 2012;12:1–20.
71. Whitney D, Levi DM. Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends Cogn Sci*. 2011;15:160–168.
72. Anderson EJ, Dakin SC, Schwarzkopf DS, Rees G, Greenwood JA. The neural correlates of crowding-induced changes in appearance. *Curr Biol*. 2012;22:1199–1206.
73. Hensch TK. Critical period regulation. *Annu Rev Neurosci*. 2004;27:549–579.

Supplementary materials for:

Visual acuity, crowding and stereo-vision are linked in children with and without amblyopia

John A. Greenwood[1,2,3,4], Vijay K. Tailor[2,5], John J. Sloper[2,5], Anita J. Simmers[6], Peter J. Bex[7], and Steven C. Dakin[1,2]

[1] *UCL Institute of Ophthalmology, University College London, London, UK*
[2] *NIHR Biomedical Research Centre for Ophthalmology at Moorfields Eye Hospital NHS Foundation Trust, London, UK*
[3] *Laboratoire Psychologie de la Perception, Université Paris Descartes, Sorbonne Paris Cité, Paris, France*
[4] *Centre National de la Recherche Scientifique, UMR 8158, Paris, France*
[5] *Strabismus and Paediatric Service, Moorfields Eye Hospital, London, UK*
[6] *Vision Sciences, Department of Life Sciences, Glasgow Caledonian University, Glasgow, UK*
[7] *Schepens Eye Research Institute, Harvard Medical School, Boston, MA, USA*

Supplementary movie files

We include four supplementary movie files to demonstrate the operation of each task in the VacMan battery. For demonstration, screen dimensions have been reduced to 640×480 pixels, though this was considerably larger in the actual task (see main text). A description of each movie follows:

Supplementary Movie 1. Two example trials of the acuity task. The first response is incorrect, causing an increase in stimulus size on the second trial. The second response is correct and the reward animation follows (though in the actual task, three correct responses were required for the animation).

Supplementary Movie 2. Two example trials of the contrast-detection task. The first response is correct and VacMan becomes dimmer on the second trial as a result.

Supplementary Movie 3. Two example trials of the crowding task. The first response is incorrect, and the centre-to-centre spacing between VacMan and the achromatic "ghost" flankers increases on the subsequent trial.

Supplementary Movie 4. Two example trials of the stereo-acuity task, with monocular "shadows" present. The presence of binocular disparity is demonstrated here by showing the signal to each eye on subsequent frames of the movie (causing the "ghost" with disparity to move back-and-forth, which was not present in the actual task). The first trial elicits a correct response, causing the binocular disparity to decrease on the subsequent trial.

Clinical details of the children

All children underwent a full orthoptic examination prior to participating in the experiments. The four tables on the following pages display the details of each child and their results in the orthoptic exam. The selection criteria for each group are described in-text, but note that there was no requirement for the presence of amblyopia in the three "clinical" groups, as we wished to assess a range of visual abilities. Though the majority do meet this criterion (either acuity worse than 0.1 logMAR or an interocular acuity difference greater than 0.1 logMAR), one of the anisometropic children (MF) and four of the strabismic children (SS, AI, JM, and HA) do not. Nonetheless, removing these children does not qualitatively alter the analyses described in-text. We also sought a mix of stereo-vision abilities for each group, and thus the observed frequencies of stereo-blindness for these groups should not be taken as indicative of the overall population (e.g. Richards, 1970, 1971).

Initials	Age (months)	Sex	Refractive error	logMAR acuity	TNO stereo.
MW	67	F	R: Plano L: Plano	R: 0.00 L: 0.00	60″
KK	70	F	R: +2.00/-1.25×180° L: +2.00/-1.25×180°	R: 0.075 L: 0.075	240″
OC	64	F	R: Plano L: Plano	R: 0.06 L: 0.02	60″
IB	85	M	R: Plano L: Plano	R: 0.00 L: 0.00	120″
BM	94	F	R: Plano L: Plano	R: 0.00 L: 0.00	60″
BR	83	F	R: Plano L: Plano	R: 0.05 L: 0.00	60″
AH	86	M	R: Plano L: Plano	R: 0.025 L: 0.025	60″
FC	75	M	R: Plano L: Plano	R: 0.04 L: 0.04	60″
KA	63	M	R: Plano L: Plano	R: 0.00 L: 0.00	120″
BD	83	F	R: Plano L: Plano	R: 0.00 L: 0.00	60″
SA	90	M	R: Plano L: Plano	R: -0.10 L: 0.00	60″
SL	75	M	R: Plano L: Plano	R: 0.025 L: 0.075	120″
JB	70	M	R: Plano L: Plano	R: 0.00 L: 0.00	60″
SA	92	F	R: Plano L: Plano	R: 0.06 L: 0.00	60″
JT	100	F	R: +2.75 DS L: +2.75/-0.50×165°	R: 0.00 L: -0.04	60″
IK	86	F	R: Plano L: Plano	R: 0.00 L: 0.00	60″
GJ	60	M	R: Plano L: Plano	R: -0.10 L: -0.10	60″
EG	60	F	R: Plano L: Plano	R: 0.00 L: 0.04	60″
GF	81	F	R: Plano L: Plano	R: 0.00 L: 0.00	60″

Supplementary Table 1. Clinical details of the 19 control children. Age is reported in months. Optical correction includes cylindrical and spherical values with appropriate axes for each eye, where R = right eye and L = left eye. logMAR acuity is similarly reported for each eye, and results of the TNO stereo-acuity test are reported in seconds of arc.

Initials	Age (months)	Sex	Refractive error	logMAR acuity	TNO stereo.
GF	102	M	R: +7.00/-1.00×180° L: +6.00/-1.00×180°	R: 0.10 L: 0.10	360″
AK	83	F	R: +1.00/-0.5×5° L: +3.50/-0.5×180°	R: 0.00 L: 0.125	60″
SA	85	F	R: -5.75/-0.5×20° L: +1.75/-1.25×20°	R: 0.25 L: 0.05	Nil
RH	101	M	R: +4.50 DS L: +5.50/-2.25×175	R: 0.025 L: 0.125	120″
AV	98	F	R: +1.25 DS L: +7.00/-1.00×180°	R: -0.22 L: 0.10	240″
AA	80	M	R: -0.25/-1×25° L: 0/-2.00×165°	R: 0.00 L: 0.10	60″
BH	72	M	R: Plano L: +4.50 DS	R: -0.10 L: 0.10	120″
CT	95	M	R: +6.50/-0.50×60° L:+1.75 DS	R: 0.10 L: -0.10	240″
JK	80	F	R:+2.50/-0.25×180° L:+6.00/-1.75×170°	R: 0.00 L: 0.10	120″
JB	86	M	R:+1.00 DS L:+3.00 DS	R: 0.00 L: 0.28	120″
NR	104	F	R:+5.50/-1.25×180° L:+1.00/-0.50×5°	R: 0.60 L: 0.00	Nil
PS	90	F	R:+5.75/-1.75×10° L:+2.75/-0.50×170°	R: 0.14 L: 0.00	120″
MF	83	F	R: -0.50/-0.50×180° L:+1.25/-0.25×180°	R: 0.00 L: 0.00	60″
CB	62	M	R:+3.25/-1.5×10° L:+2.00/-0.75×180°	R: 0.80 L: 0.10	Nil
LC	108	M	R:+5.50 DS L:+3.75 DS	R: 0.12 L: -0.10	60″
SS	56	F	R:+3.25/-3.50×95° L:+1.75/-1.00×95°	R: 0.20 L: 0.00	160″

Supplementary Table 2. Clinical details of the 16 anisometropic children. Values reported are in the same format as Supplementary Table 1.

Initials	Age (months)	Sex	Ocular alignment (with Rx)	Refractive error	logMAR acuity	TNO stereo.
MO	80	F	n: L Alt. SOT 16Δ d: L Alt. SOT 6Δ	R: +5.50/-3.00×180° L: +5.75/-2.50×165°	R: 0.275 L: 0.45	Nil
SS	92	F	n: straight d: R XOT 18Δ	R: Plano L: Plano	R: 0.00 L: 0.00	60″
NJ	94	M	n: L XOT 10Δ d: L XOT 8Δ L/R 10Δ	R: +3.75/-0.25×10° L: +4.50/-0.25×70°	R: 0.10 L: 0.15	Nil
EG	60	F	n: straight d: R XOT 20Δ	R: +1.00/-0.50×90° L: +1.50 DS	R: 0.15 L: 0.00	240″
GD	92	F	n: L SOT 45Δ d: L SOT 40Δ	R: +5.75/-1.25×175° L: +6.50/-1.25×180°	R: 0.15 L: 0.75	Nil
DS	94	M	n: R SOT 8Δ R/L 6Δ d: R SOT 6Δ R/L 11Δ	R: Plano L: Plano	R: 0.20 L: 0.00	Nil
HD	83	M	n: R SOT 10Δ d: R SOT 6Δ	R: +8.50/-0.50×155° L: +8.25/-2.00×45°	R: 0.36 L: 0.10	Nil
JB	64	M	L SOT 20Δ	R: +4.00/+0.50×180° L: +4.75/-0.75×175°	R: 0.06 L: 0.46	Nil
NO	96	M	L SOT <10Δ	R: +1.50 DS L: +2.25 DS	R: -0.14 L: 0.06	60″
OC	105	M	R SOT 12Δ	R: +7.00/-1.00×20° L: +7.50/-1.50×170°	R: 0.54 L: 0.10	Nil
RS	75	M	L SOT 40Δ	R: +8.25/-1.00×30° L: +8.25/-1.50×160°	R: 0.10 L: 0.76	Nil
LH	60	F	L SOT <10Δ	R: +6.25/-0.75×180° L: +6.50 DS	R: 0.00 L: 0.18	Nil
AI	53	F	L SOT <10Δ	R: +5.50/-0.50×180° L: +5.75/-0.25×30°	R: 0.00 L: 0.06	Nil
TP	95	F	R SOT 10Δ	R: +6.50/-2.50×140° L: +6.50/-2.75×170°	R: 0.40 L: 0.10	Nil
CA	97	F	n: L SOT 45Δ d: L SOT 25Δ	R: +2.50 DS L: +2.25 DS	R: 0.10 L: 0.15	Nil
JM	58	M	Alt. R SOT 15Δ	R: +1.50 DS L: +1.25 DS	R: 0.06 L: 0.02	Nil
HA	61	F	n: Alt. R SOT 20Δ d: Alt. R SOT 10Δ	R: +4.50 DS L: +5.00 DS	R: 0.02 L: 0.04	60″
HD	61	F	R SOT 20Δ	R: +6.00 DS L: +6.00 DS	R: 0.26 L: 0.00	Nil

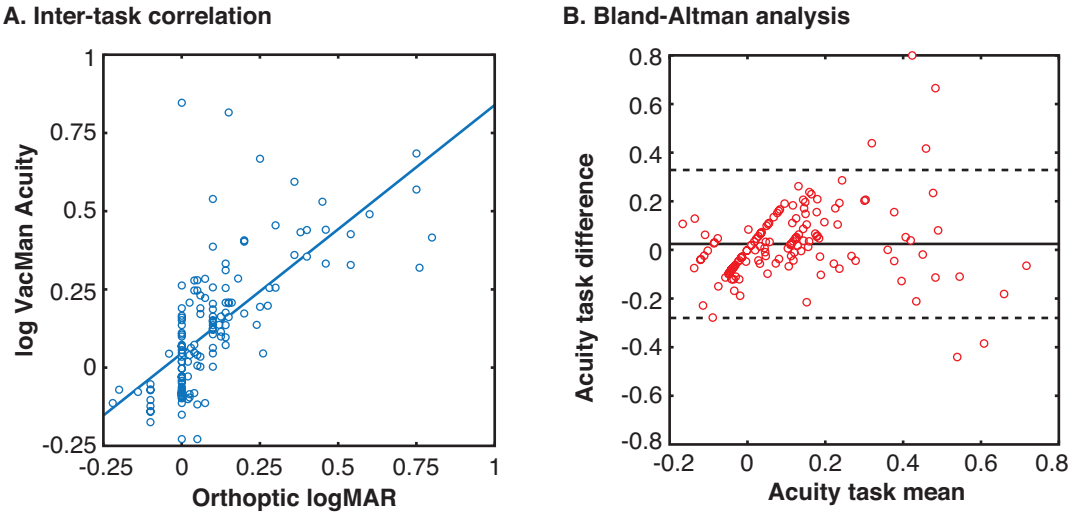
Supplementary Table 3. Clinical details of the 18 strabismic children. The "ocular alignment" column reports the outcome of both near (n) and distance (d) prism cover tests. Here, SOT = esotropia, XOT = exotropia, Alt. = alternating, L/R = left over right, and R/L = right over left. The degree of deviation is shown in prism dioptres and the amblyopic eye is denoted. All other values are in the same format as the preceding tables.

Initials	Age (months)	Sex	Ocular alignment (with Rx)	Refractive error	logMAR acuity	TNO stereo.
JP	86	M	R SOT 10Δ	R: -9.75/-1.50×75° L -3.25/-1.75×125°	R: 0.75 L: 0.05	Nil
SH	95	M	L SOT <10Δ	R: +3.25 DS L: +6.75 DS	R: -0.10 L: 0.05	120″
AW	65	M	n: R SOT 25Δ d: R SOT 20Δ	R: +7.25/-1.50×20° L: +6.25/-1.50×155°	R: 0.25 L: 0.05	Nil
RM	70	M	L SOT 25Δ L/R 12Δ	R: +2.50/-0.50×180° L: +4.00/-1.75×65°	R: 0.05 L: 0.125	Nil
GE	75	F	n: R SOT 16Δ d: R SOT 12Δ	R: +4.50/-0.5×160° L: +0.75 DS	R: 0.20 L: 0.00	Nil
BW	82	M	L SOT 30Δ	R: +3.00/-1.00×100° L: +6.00/-1.50×80°	R: -0.20 L: 0.14	Nil
AJ	96	M	L SOT <10Δ	R: 0.00/-0.50×5° L: +4.00/-1.25×175°	R: 0.00 L: 0.16	240″
MM	59	F	R SOT 30Δ	R: +6.25/-1.25×115° L: +5.25/-0.75×50°	R: 0.30 L: 0.10	Nil
AW	66	M	R SOT 35Δ	R: +7.00/-0.75×180° L: +6.00/-1.00×180°	R: 0.38 L: 0.10	Nil
EO	62	F	L SOT 10Δ	R: +4.5 DS L: +6.5 DS	R: 0.10 L: 0.36	Nil
SG	104	F	R SOT <10Δ	R: +5.50/-4.00×10° L: +3.00/-1.50×160°	R: 0.14 L: -0.10	60″
DB	84	F	L SOT <10Δ	R: +2.75 DS L: +4.25 DS	R: 0.00 L: 0.30	360″
EP	80	F	L SOT <10Δ	R: +1.5/-0.5×180° L: +6.25/-2.25×10°	R: 0.00 L: 0.40	Nil
TH	107	M	L SOT 12Δ L/R 9Δ	R: +1.75/-0.5×20° L: +3.00/-1.5×170°	R: 0.00 L: 0.54	Nil
HB	70	F	L SOT <10Δ	R: +4.25/-0.50×150° L: +5.25/-0.75×170°	R: 0.14 L: 0.46	Nil
NO	77	F	R SOT 18Δ	R: +4.5/-0.25×180° L: +3.5/-0.50×180°	R: 0.14 L: 0.03	Nil
NT	73	F	L SOT <10Δ	R: +4.25 DS L: +6.25 DS	R: 0.04 L: 0.14	450″
DB	93	M	n: R SOT 25Δ d: R SOT 14Δ	R: +4.00 DS L: +2.75/+0.75×180°	R: 0.24 L: 0.00	Nil
JD	77	F	SOT 30Δ	R: +3.50/-0.75×5° L: +5.00/+3.50×175°	R: 0.14 L: 0.14	Nil

Supplementary Table 4. Clinical details of the 19 mixed anisometropic and strabismic children. All values are reported in the same format as Supplementary Table 3.

Inter-task correlation and agreement

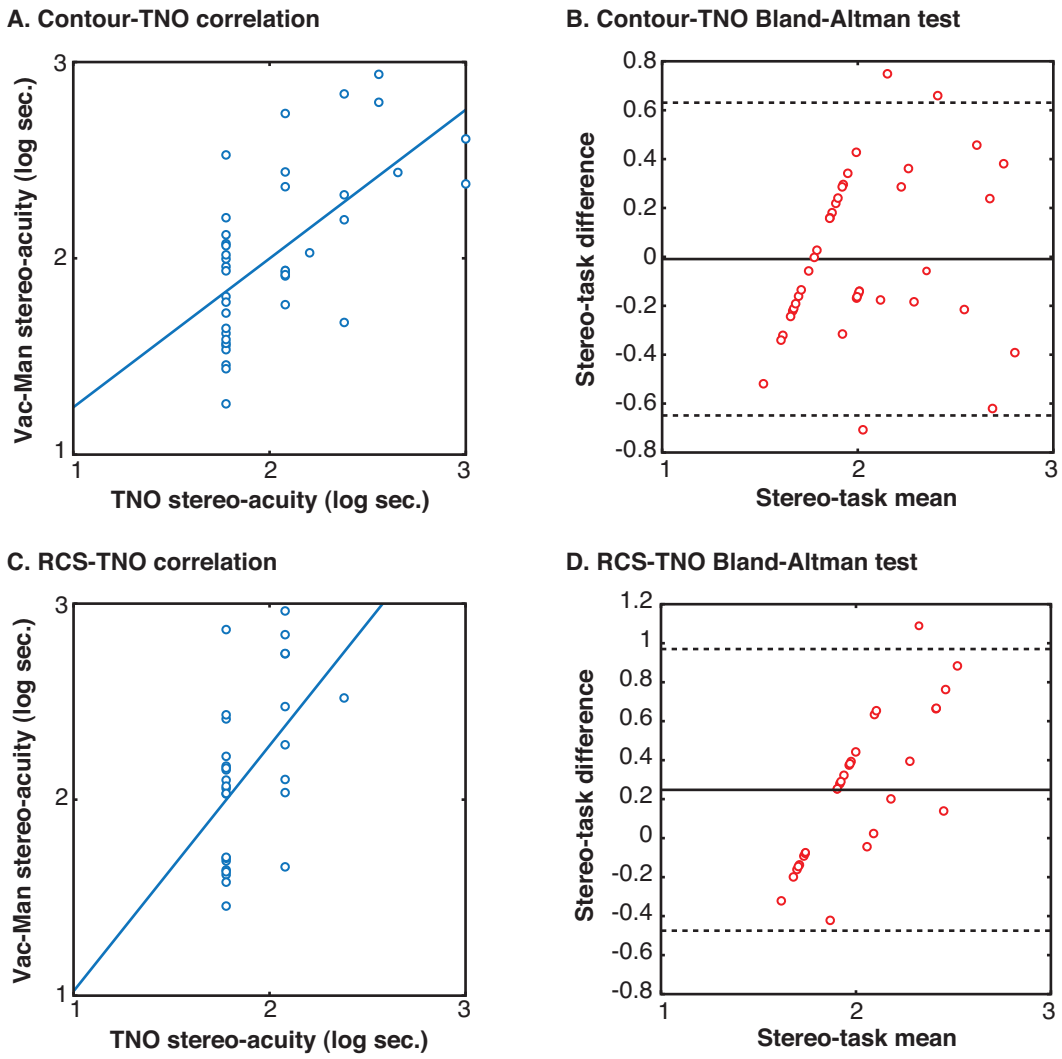
Because all children underwent a full orthoptic examination prior to their inclusion in the study, we can compare values obtained using our *Vac-Man* procedures and those obtained using more established methods. In the clinic, logMAR acuity was measured with the Thompson V2000. When these values are compared with our acuity thresholds (in equivalent units of log minutes of arc), a strong correlation is observed: $r(142) = 0.68$, $p<.001$ (Supplementary Figure 1A). A Bland-Altman analysis (Bland & Altman, 1986) was conducted to examine the level of agreement between these values. This involves a comparison between the mean of acuity values on each task as a function of the difference between these values, as plotted in Supplementary Figure 1b. This shows an acceptable level of agreement between the two tasks: 95% of our acuity values were between ±0.31 log min. (dashed lines), around a mean difference of 0.02 log minutes.



Supplementary Figure 1. A. Correlation between our VacMan acuity thresholds and logMAR values obtained during orthoptic examination. Our values (in log minutes of arc) correlate strongly with orthoptic values. **B.** Bland-Altman test between VacMan acuity thresholds and the orthoptic logMAR values. The x-axis plots the average of the orthoptic logMAR values and log VacMan Acuity values, while the y-axis plots the difference between the two values. Dashed lines indicate the 95% confidence limits on the differences between values, which indicate that the differences largely fall within a range of ±0.31 log minutes around a mean difference (solid line) of 0.02 log minutes.

Stereo-acuity was also assessed during the orthoptic examinations, using the TNO circles test. For comparison with our Vac-Man measures, we converted these values into log seconds of arc. If we examine only those with measurable stereo-thresholds on the contour-based Vac-Man stereo task, there is a strong correlation between these two

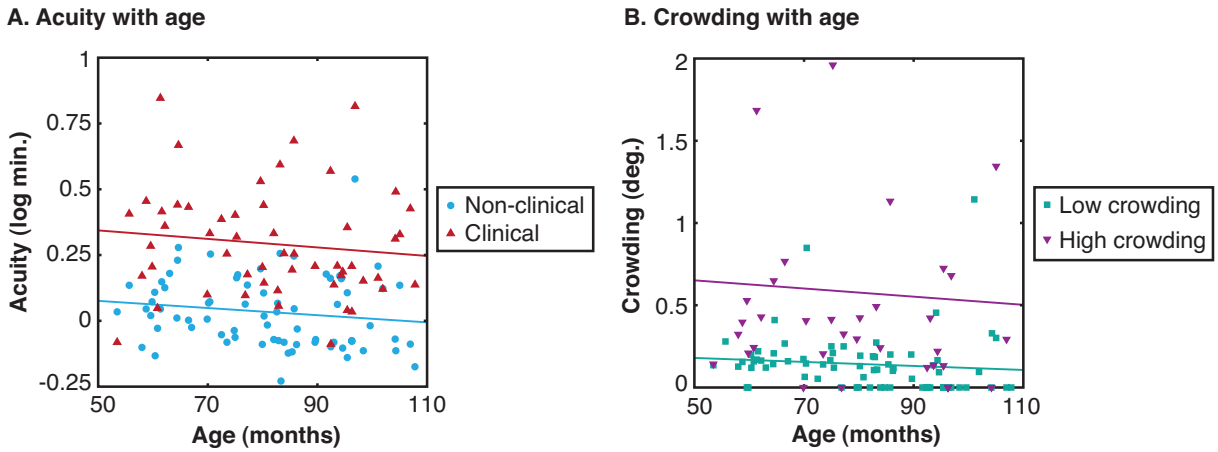
measures ($r(39)=0.63$ $p<0.001$), as plotted in Supplementary Figure 2A. When submitted to a Bland-Altman analysis, as in Supplementary Figure 2B, these values show modest agreement, with 95% of values falling within ±0.64 log seconds around a mean of -0.01 log seconds. Similarly, for the random-check stereogram (RCS) version of the task, there was a strong correlation with TNO performance ($r(28)=0.49$ $p=0.01$), as plotted in Supplementary Figure 2C, despite the lower number of children with measurable thresholds in this task. The Bland-Altman analysis for these tasks, plotted in Supplementary Figure 2D, reveals some degree of bias – Vac-Man RCS thresholds were consistently higher than those obtained in the TNO, with a mean difference of 0.25 log seconds. The spread of error is nonetheless similar, with 95% of values contained within ±0.72 log seconds. The discrepancy between these stereo-measures is likely due to our Vac-Man task being conducted at a distance of 3m, compared to the arm's length distance of the TNO test. Although this greater distance has the advantage of minimising potential motion parallax cues (Ono, Rivest, & Ono, 1986), it also increases the likelihood that individual checks in the display will fall below acuity limits and fail to be resolved. The same is not true for the contour-based task where lower spatial frequency information is also available. Future experiments using the Vac-Man stereo-task may nonetheless benefit from closer viewing distances.



Supplementary Figure 2. A. Correlation between our Vac-Man stereo-acuity thresholds, obtained with monocular "shadows" present, and values obtained from the TNO circles test during orthoptic examination. Both values (in log seconds of arc) correlate strongly. **B.** Bland-Altman test between our contour-based stereo-acuity thresholds and the TNO values. The x-axis plots the average of the orthoptic two values, while the y-axis plots the difference. Dashed lines indicate the 95% confidence limits on the differences between values. **C**. Correlation between our Vac-Man stereo-acuity thresholds, obtained in the random-check stereogram (RCS) version of the task, and TNO values, plotted as in panel A. **D**. Bland-Altman test between our random-check stereogram Vac-Man test and the TNO circles test, plotted as in panel B.

The effect of age

We include in our sample of children an age range from 54-107 months, with a mean of 80.7 months. To examine whether this age range contains variation in abilities, we examined the effect of age on acuity and crowding. Prior studies have shown that acuity is typically adult-like by approximately 6 years of age (Atkinson, Anker, Evans, & McIntyre, 1987; Ellemberg, Lewis, Hong Liu, & Maurer, 1999; Pan et al., 2009; Jeon, Hamid, Maurer, & Lewis, 2010), approximately midway through our age-range. The improvement over time for amblyopic children will obviously be confounded by the age at which treatment regimes begin, and so we separate amblyopic eyes from the fellow eyes and those of controls. Supplementary Figure 3A plots the relationship between acuity and age (in months) for these two groups of values. The "non-amblyopic" group includes averaged values for each control subject, and the fellow eye of all our amblyopic children. Computed in this way, there is no relationship between acuity and age over this range, $r(70)=-.15$, $p=.22$. Similarly, there is no improvement in acuity for the amblyopic eyes of children in the three clinical groups, $r(51)=-.12$, $p=.38$.



Supplementary Figure 3. Age effects on acuity and crowding. **A.** The effect of age (in months) on acuity values (in log minutes of arc). Here, "non-amblyopic" values (blue circles) represent the average of both eyes (for control subjects) or the fellow eye of children in the three clinical groups. The "amblyopic" values (red triangles) represent the amblyopic eye of children in all three clinical groups. There is no relationship with age for either group. **B.** The effect of age on crowding (in degrees). Here the "low crowding" values (green squares) represent the average of both eyes (for control subjects) or the fellow eye of children in the three clinical groups, while the "high crowding" values are the amblyopic eyes of children in the strabismic and mixed groups. There is no relationship for either dataset.

In contrast with the early development of acuity, elevated foveal crowding has been observed in "normally" developing children as late as 11 years (Atkinson & Braddick,

1983; Atkinson et al., 1987; Jeon et al., 2010). However, it is possible that age may improve crowding in the amblyopic group due to the progression of treatment in this time. To examine this, data was separated into those with "high crowding", i.e. the amblyopic eye of children in the strabismic and mixed groups, with comparison to those with "low crowding" – an average of foveal values in control children, and the fellow eye of all amblyopes. When computed this way, as plotted in Supplementary Figure 3B, there is no relationship between age and crowding (in degrees) for either of the "low crowding" ($r(70)=-0.10$, $p=.41$) or "high crowding" ($r(35)=-0.06$, $p=.72$) datasets.

Correlations between stereo-acuity thresholds and the spatial tasks

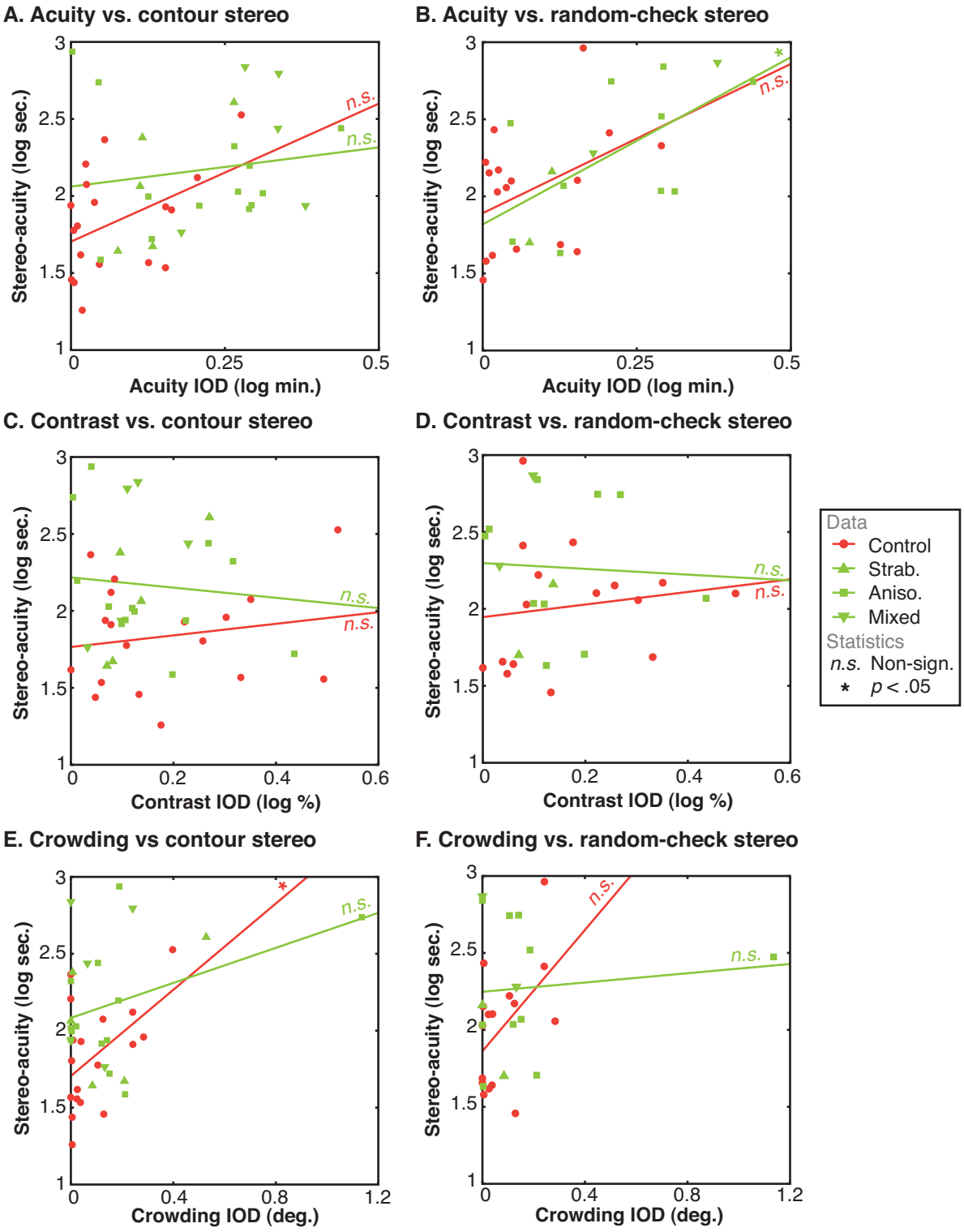
In the main-text, Figure 4 displays the stereo-acuity thresholds for those with measurable stereo-abilities. To examine the relationship between stereo-vision and interocular differences (IODs) in each of the three spatial tasks (acuity, contrast-detection, and crowding), we divided children into "stereo" and "non-stereo" categories based on their ability to achieve measureable stereo-acuity thresholds (i.e. below 1000" of arc) on *either* of the two stereo-tests. As outlined in-text, we considered this to be the conservative method of comparison rather than assigning arbitrary values to the "non-stereo" children in each group. In this section we examine the correlation between the stereo-thresholds achieved by children in the "stereo" category and their IODs in the three spatial tasks.

Because it is the existence of any interocular difference that may disrupt stereo-acuity (Goodwin & Romano, 1985), the direction of interocular differences is unimportant. We thus took the absolute interocular difference in the three spatial tasks. This does not affect data in the three clinical groups considerably, but does alter the distribution of control values. The relationship between contour-based stereo-thresholds and absolute interocular differences in acuity are plotted in Supplementary Figure 4A. Overall, this relationship is significant, with a correlation of $r(39)=0.41$, $p=0.008$. Because the selection of only those children with measureable stereo-thresholds reduces group numbers substantially (particularly for the strabismic and mixed groups), here we combine the three clinical groups into a single category. When combined in this way, there was a marginally non-significant correlation with interocular acuity differences in the control

group (*r(16)*=0.44, *p*=0.068), and a non-significant relationship for the clinical group (*r(21)*=0.15, *p*=0.51). Both groups can however be seen to follow roughly the same pattern, albeit noisily. Similarly, Supplementary Figure 4B plots the relationship between the random-check stereogram version of our task and IODs in acuity. Overall, this relationship was highly significant, with *r*(28)=0.57, *p*=0.001. When divided by category, there was a non-significant correlation for the control group (*r(14)*=0.34, *p*=0.20), though this data can be seen to follow the same trend as the clinical group for whom the relationship was significant (*r(12)*=0.63, *p*=0.016).

Supplementary Figure 4C plots the relationship between contrast-detection thresholds and contour-based stereo. Overall there is no significant correlation between the two (*r*(39)=-0.03, *p*=0.84), which is true for the individual groups as well (control: *r(16)*=0.17, *p*=0.49; clinical: *r(21)*=-0.09, *p*=0.70). The same is true for random-check stereo-acuities, as plotted in Supplementary Figure 4D. There is no significant correlation either overall (*r*(28)=0.01, *p*=0.94), or for the separate groups (control: *r(14)*=0.14, *p*=0.60; clinical: *r(12)*=0.05, *p*=0.87).

For contour-based stereo, the relationship with crowding is similar to that of acuity. The two are plotted in Supplementary Figure 4E, and overall this relationship is highly significant (*r*(39)=0.39, *p*=0.011). When broken down, this relationship is significant for the control group (*r(16)*=0.49, *p*=0.038) but not for the clinical group (*r(21)*=0.34, *p*=0.109). In contrast, there was no significant relationship between random-check stereo-acuities and crowding for the data as a whole (*r*(28)=0.23, *p*=0.22), nor for the clinical group separately (*r(12)*=0.10, *p*=0.73), though the control children show a marginally non-significant correlation (*r(14)*=0.49, *p*=0.06).



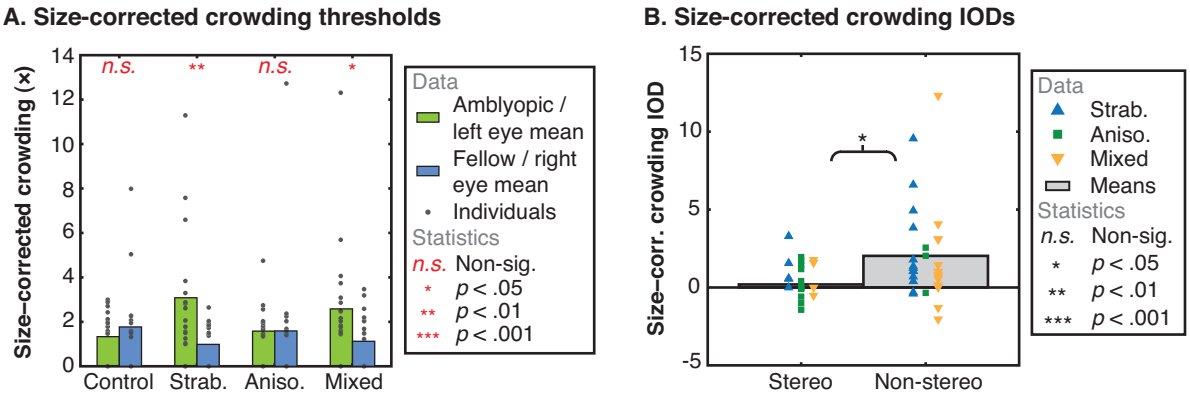
Supplementary Figure 4. Correlations between the three spatial tasks and stereo-acuity thresholds, for observers with measurable stereo-acuity. **A.** The relationship between contour-based stereo-acuity thresholds (in log seconds of arc) and interocular differences (IODs) in acuity (in log minutes). Children are divided into "control" and "clinical" categories, though separate symbols show whether clinical children are derived from the strabismic (upright triangles), anisometropic (squares), or mixed (inverted triangles) groups. The correlation for each group is shown at the right-hand side of the line of best fit. **B.** Random-check stereo-acuity thresholds, plotted against IODs for acuity, as in panel A. **C.** Contour-based stereo and IODs for contrast detection, the latter plotted in log Weber percentage units. **D.** Random-check stereo-acuity thresholds and contrast-detection IODs. **E.** IODs for crowding (in degrees of visual arc), plotted against contour-based stereo-acuity thresholds. **F.** Crowding IODs and random-check stereo-acuity thresholds.

Together, we see a similar pattern in these results as when children are divided into "stereo" and "non-stereo" categories (as in Figure 5 in-text). Namely, those with poor stereo-acuity have large interocular differences in both acuity and crowding. The correlation between IODs in acuity and stereo-thresholds has been observed previously in adults (Weakley, 2001; McKee, Levi, & Movshon, 2003; Levi, McKee, & Movshon, 2011); here we demonstrate that this is true for both contour-based and random-check stereo stimuli. Differences emerge with crowding however, where we demonstrate a strong relationship between contour-based stereo-acuity and crowding, while the relationship between random-check stereo-acuity and crowding is non-significant. This could be due to the stronger demands placed on acuity by the random-check stereograms – the individual checks in these stimuli must be resolved in order to detect the interocular correlations. The relationship with acuity may dominate these thresholds as a result, whereas the additional low spatial frequency information in the contour-based stereograms may allow the relationship with crowding to emerge.

Size-corrected crowding values

In the "normal" visual system, the extent of foveal crowding in adults has been found to follow the size of the stimuli used in testing (Levi, Klein, & Hariharan, 2002), while in the periphery (Levi, Hariharan, & Klein, 2002b; Tripathy & Cavanagh, 2002) and the amblyopic fovea (Levi, Hariharan, & Klein, 2002a), crowding is largely size-invariant. In our study, the large variation in acuity within our population necessitated that stimuli in the remaining tasks was scaled to clearly visible sizes to ensure visibility. However, this variation in stimulus size raises the possibility that the elevations in foveal crowding seen in some groups could simply reflect the variation in stimulus size (and the corresponding variation in crowding). To examine this possibility, our measured extents of crowding (in degrees of visual angle) were divided by the size of stimuli used in this task (also in degrees, which corresponds to values 2.5× acuity thresholds). The resulting values represent the spatial extent of foveal crowding expressed in multiples of stimulus size. Note that this is equivalent to expressing the extent of crowding in units of "letter widths", as is sometimes performed for crowding experiments.

Supplementary Figure 5A plots the spatial extent of this size-corrected foveal crowding in each eye for all children in our study. The same pattern of results as in Figure 2C is clearly evident. There is no interocular difference between the degree of crowding in either of the control ($t(18)$=-1.16, p=.26) or anisometropic ($t(15)$=-0.02, p=.99) groups, while a clear interocular difference is evident in the strabismic group ($t(17)$=3.38, p=.004) and a marginally significant difference exists for the mixed group ($t(18)$=2.10, p=.049). On top of this, there is still a clear degree of crowding in both eyes of controls and the fellow eyes of all clinical groups (slightly below two times the stimulus size in each case), with a combined analysis showing a significant difference ($t(90)$=5.63, p<.001). This is consistent with prior studies (Jeon et al., 2010).



Supplementary Figure 5. Size-corrected crowding values. **A.** When crowding values are expressed in multiples of acuity thresholds, the degree of foveal crowding remains significant in all eyes, as do interocular differences for the strabismic and mixed groups (data as in Figure 2 in-text). **B.** Interocular crowding differences are also larger for "non-stereo" children than for those with measurable stereo-vision (data plotted as in Figure 5 in-text).

It is also possible that the higher levels of crowding observed in our non-stereo subjects (Figure 5C of the main text) reflects the potential size-tuning of foveal crowding. To examine whether the larger IODs persisted after size correction, we split our observers into "stereo" and "non-stereo" categories, as in the main text, and plot size-corrected crowding extents in Supplementary Figure 5B. Both the pattern of individual observers, and the large group difference on average are maintained, which remains significant ($t(51)$=-2.39, p=.021). The relationship between stereo-vision and crowding is thus independent of stimulus size.

Supplementary References

Atkinson, J., Anker, S., Evans, C., & McIntyre, A. (1987). The Cambridge Crowding Cards for preschool visual acuity testing. *Transactions of the Sixth International Orthoptic Congress*, 482-486.

Atkinson, J., & Braddick, O. J. (1983). Assessment of visual acuity in infancy and early childhood. *Acta Ophthalmologica, 157 Suppl*, 18-26.

Bland, M. J., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet, 327*(8476), 307-310.

Ellemberg, D., Lewis, T. L., Hong Liu, C., & Maurer, D. (1999). Development of spatial and temporal vision during childhood. *Vision Research, 39*(14), 2325-2333.

Goodwin, R. T., & Romano, P. E. (1985). Stereoacuity degradation by experimental and real monocular and binocular amblyopia. *Investigative Ophthalmology & Visual Science, 26*(7), 917-923.

Jeon, S. T., Hamid, J., Maurer, D., & Lewis, T. L. (2010). Developmental changes during childhood in single-letter acuity and its crowding by surrounding contours. *Journal of Experimental Child Psychology, 107*(4), 423-437.

Levi, D. M., Hariharan, S., & Klein, S. A. (2002a). Suppressive and facilitatory spatial interactions in amblyopic vision. *Vision Research, 42*(11), 1379-1394.

Levi, D. M., Hariharan, S., & Klein, S. A. (2002b). Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision, 2*, 167-177.

Levi, D. M., Klein, S. A., & Hariharan, S. (2002). Suppressive and facilitatory spatial interactions in foveal vision: Foveal crowding is simple contrast masking. *Journal of Vision, 2*(2), 140-166.

Levi, D. M., McKee, S. P., & Movshon, J. A. (2011). Visual deficits in anisometropia. *Vision Research, 51*(1), 48-57.

McKee, S. P., Levi, D. M., & Movshon, J. A. (2003). The pattern of visual deficits in amblyopia. *Journal of Vision, 3*(5), 380-405.

Ono, M. E., Rivest, J., & Ono, H. (1986). Depth perception as a function of motion parallax and absolute-distance information. *Journal of Experimental Psychology: Human Perception and Performance, 12*(3), 331-337.

Pan, Y., Tarczy-Hornoch, K., Cotter, S. A., Wen, G., Borchert, M. S., Azen, S. P., et al. (2009). Visual acuity norms in pre-school children: The multi-ethnic pediatric eye disease study. *Optometry & Vision Science, 86*(6), 607-612.

Richards, W. (1970). Stereopsis and stereoblindness. *Experimental Brain Research, 10*(4), 380-388.

Richards, W. (1971). Anomalous stereoscopic depth perception. *Journal of the Optical Society of America, 61*(3), 410-414.

Tripathy, S. P., & Cavanagh, P. (2002). The extent of crowding in peripheral vision does not scale with target size. *Vision Research, 42*(20), 2357-2369.

Weakley, D. R. (2001). The association between nonstrabismic anisometropia, amblyopia, and subnormal binocularity. *Ophthalmology, 108*(1), 163-171.