Current Biology in press

Young children do not integrate visual and haptic form information

Monica Gori[1,2], Michela Del Viva[3,4], Giulio Sandini[1,2] and David C. Burr[3,5]

1. Istituto Italiano di Tecnologia, via Morego 30, 16163 Genoa, Italy
2. Dipartimento di Informatica Sistemistica e Telematica, Via all'Opera Pia 13, 16145 Genoa, Italy
3. Dipartimento di Psicologia, Università Degli Studi di Firenze, Via S. Nicolò 89, Florence, Italy.
4. Istituto di Neuroscienze del CNR, Via Moruzzi 1, Pisa, Italy.
5. Department of Psychology, University of Western Australia, Perth WA, Australia.

Key words: Perception – development – vision – touch

ABSTRACT

Several studies have shown that adults integrate visual and haptic information (and information from other modalities) in a statistically optimal fashion, weighting each sense according to its reliability [1, 2]. When does this capacity for cross-modal integration develop? Here we show that prior to eight years of age, integration of visual and haptic spatial information is far from optimal, with either vision or touch dominating totally, even in conditions where the dominant sense is far less precise than the other (assessed by discrimination thresholds). For size discrimination, haptic information dominates in determining both perceived size and discrimination thresholds, while for orientation discrimination vision dominates. By eight-ten years, the integration becomes statistically optimal, like adults. We suggest that during development, perceptual systems require constant recalibration, for which cross-sensory comparison is important. Using one sense to calibrate the other precludes useful combination of the two sources.

RESULTS

We measured visual-haptic integration of two aspects of form perception in young (5-10 year-old) children: size and orientation discrimination. The size discrimination task was a low-technology, child-friendly adaptation of Ernst and Banks' [1] technique, where subjects were required to discriminate the height of physical blocks on the basis of visual, haptic or visuo-haptic information (see Fig. 1A). As this technique differed in some respects to the more standard virtual reality techniques, we first validated it with adults to demonstrate that optimal cross-modal integration did occur under these conditions [see also 3]. The results (reported in Sup. Mat., along with detailed illustration and description of the stimuli) were very similar to those obtained by Ernst and Banks [1]: with various levels of visual stimulus degradation (via image blur), perceived size of visual-haptic stimuli followed closely the maximum likelihood estimate (MLE) predictions and, most importantly, the thresholds for dual-modality presentation were lower than either visual or haptic thresholds, the main signature for cross-modal integration.

We then proceeded to measure haptic, visual and bimodal visuo-haptic size discrimination in 5-10 year-old children. Children were presented two successive stimuli and asked to judge in two-alterative forced choice which was the taller (guessing if unsure). For the visual and haptic trials, one stimulus (randomly first or second) was the *standard*, always 55 mm high, and the other the *probe*, of variable height between 48 and 62 mm. The proportion of trials where the probe was judged taller than the standard was computed for each probe height, and were well fitted by cumulative Gaussian functions (Figs. 1B&C). The mean of the fitted Gaussian estimates the point of subjective equality (PSE), near zero for all uni-modal conditions, showing there was no bias in perceived size of probes and tests. The standard deviation (inverse slope) of the curves estimates discrimination thresholds. In these two example subjects, the steeper curves for the visual discriminations shows that visual thresholds were slightly lower than haptic thresholds, and that for both senses, thresholds for the 10 year-old were lower than for the 5 year-old. The red and green symbols of Fig. 1D show how average haptic and visual thresholds varied with age. For both senses, thresholds improved by about 30% over this age range, and at all ages haptic thresholds are about twice visual thresholds.

We also measured size discrimination in a dual-modality condition, where both visual and haptic information were provided, *in conflict*: the standard now comprised visual and haptic blocks of different heights, the visual block 55+Δ mm and the haptic block 55–Δ mm (Δ = 0 or ±3 mm). In the probe the visual and haptic stimuli varied congruently, again between 48 and 62 mm. Despite the visuo-haptic conflict of the standard, the blocks appeared as one single stimulus; no adult or child ever noticed the conflict, even when specifically questioned. Fig. 2 shows for four children sample psychometric functions for the dual-modality measurements.. The pattern of results for the 10 year-old (Fig. 2A) was very much like those for the adult (Sup. Mat.): negative values of Δ caused the curves to shift leftwards, positive values caused it to shift rightwards. That is to say the curves followed the visual standard, suggesting that visual information was dominating the match. This is consistent with the MLE model (indicated by colour-coded arrows below abscissae) , that computes a weighted average of the visual and haptic stimuli, with weights inversely related to precision threshold measured separately each single-modality: the visual judgment was more precise, and should therefore dominate (see Sup. Mat. for details of calculation). For the 5 year-old (Fig. 2B), however, the results were dramatically different: the psychometric functions for the dual-modality presentation shifted in the direction opposite to that of the 10 year-old, following the bias of the haptic stimulus. The MLE predictions are similar for both the 10 and 5 year-olds, as for both children visual thresholds were much lower than and haptic thresholds, so the visual stimuli should dominate. This does occur for the 10 year old, but for the 5 year old the reverse holds, with the haptic standard dominating the match. These results were representative of all the children we tested (shown in Sup. Mat. and summarized in Fig. 3).

The dark blue symbols of Fig. 1D show how the average dual-modality thresholds vary with age, to examine multi-sensory improvement in performance (the signature of cross-modal integration). The light-blue symbols show the thresholds predicted from the MLE model [1 and eq. 3 of Sup. Mat.]. The predicted improvement is strongest in conditions where the single-modality thresholds are most similar, such as the visually blurred condition for adults (right hand point: details in Sup. Mat.). Here the dual-modality thresholds were significantly lower than visual thresholds ($t(2) = 9.76$, p = 0.005 (one-tailed)), and statistically indistinguishable from the predicted values ($t(2) = 0.61$, $p = 0.60$ (two-tailed)). For the unblurred condition for adults and older children, the cross-modal thresholds were close to the best single-modality condition (vision), as was the MLE prediction. For the five year-olds, however, the dual-modality thresholds were as high as the haptic thresholds ($t(7) = 1.13$, $p = 0.28$ (two-tailed)), not only much higher than the MLE predictions ($t(7) = 4.76$, $p<0.05$ (one-tailed)), but twice the best single-modality (visual) thresholds ($t(7) = 4.07$, $p<0.05$ (one-tailed)). This reinforces the PSE data in showing that these young children do not integrate cross-modally in a way that benefits perceptual discrimination.

In order to ascertain whether the haptic dominance was a general phenomenon, or specific to size judgments, we repeated the series of experiments with another spatial task, orientation discrimination; a very basic visual task which could in principle be computed by neural hardware of primary visual cortex [4]. The procedure was similar to the size discrimination task, again using a simple, low-technology technique (Fig. 1E; see Sup. Mat. for full details and adult validation). Figs F&G show examples of psychometric functions for visual and haptic discriminations. As for the size judgments, the PSEs are near zero, and under these conditions (oblique standard), the visual and haptic thresholds of both the 10 and 5 year-old were similar to each other. Figs. 1F&G show sample psychometric functions for the dual-modality measurements for a 5 and 8 year-old child. As with the size judgments, the pattern of results for the 8 year-old was very much like those for the adult, with the functions of the three different conflicts (fig.2C) falling very much together, as predicted from the single modality thresholds by the MLE model (arrows under abscissae). Again, however, the pattern of results for the 5 year-old was quite different (fig.2D). Although the MLE model predicts similar curves for the three conflict conditions, the psychometric functions followed very closely the visual standards (indicated by the arrows above the graphs), the exact opposite pattern to that observed for size discrimination.

Fig. 1H shows how average thresholds varied with age. As with size discrimination, uni-modal thresholds decreased with age, but more so, a factor of four for haptic and five for visual thresholds over the age range. The dual-modality thresholds and MLE-predictions are shown by the dark- and light-blue symbols. For adults, dual-modality thresholds were lower than visual thresholds (marginally significant: $t(2) = 2.59$, $p = 0.06$ (one-tailed)), and statistically indistinguishable from the predicted values ($t(2) = 0.71$, $p = 0.54$ (two-tailed)), while for five year-olds they remain significantly higher than the predictions ($t(19) = 2.60$, $p=0.01$ (one-tailed)). Again the thresholds reinforce the PSE data in showing that these young children do not integrate cross-modally in a way that benefits perceptual discrimination.

To examine further the development of visuo-haptic integration, Fig. 3 reports PSEs for all children of all ages for the three conflict conditions, for both size and orientation discriminations, plotted as a function of the MLE predictions from single-modality discrimination-thresholds (eq. 1 & 2 Sup. Mat.). If the MLE prediction held, the data should fall along the black dotted equality line. For adults, this was clearly so, for both size and orientation. However, at 5 years of age the story was quite different. For the size discriminations, not only do the measured PSEs not follow the MLE predictions, but they run in the orthogonal direction. The data for the six-year-olds similarly do not follow the prediction, but there is a tendency for the data to be more scattered rather than ordered orthogonal to the prediction line. By eight years of age the data begin to follow the prediction, and by ten fall along it well, similar to the adult pattern of results. For orientation judgments, the MLE model predicts less variation with Δ (as the visual and haptic thresholds were more similar): but the 5 year-old data vary over the whole range, as they follow the orientation of the visual standards, and by eight years of age the data begin to follow the prediction, and nearly perfect for the adults.

Fig. 4 summarizes how visuo-haptic integration develops with age. Fig. 4A plots the amount of variance in PSEs explained by three models, MLE, visual and haptic dominance. For adults the MLE model accounts well for both size and orientation matches, with R^2 always in excess of 0.7. Visual dominance also explains well the unblurred data (as is to be expected), but when all three blur conditions are considered only the MLE model was better than the mean. For five year-olds, however, only the haptic-dominance model was better than the mean for size judgments, and vision-dominance for orientation judgments. For both tasks, the MLE predictions improved with age to become similar to adults at 8 or 10 years. Fig. 4B tells a similar story, plotting the development of theoretical and observed visual and haptic weights: violet symbols show the theoretical MLE-predicted weights (eq. 2 of Sup. Mat.), and the black symbols the actual weights that were applied for the judgments, calculated from the PSE vs conflict functions (eq 6 of Sup. Mat.). For both size and orientation, the theoretical haptic weights are fairly constant over age, around 0.2 – 0.3 (implying visual weights of 0.7 – 0.8) for size and 0.4 – 0.5 (visual weights of 0.5 – 0.6) for orientation. However, the haptic weights necessary to predict the 5 year-old PSE data are 0.6 – 0.8, far, far greater than the prediction, implying that these young children give far more weight to touch for size judgments than is optimal, as predicted by their discrimination precision. For orientation the reverse holds. Visual weights necessary to predict the 5 year-old PSE data were near unity, implying a total visual dominance. As distinct from size judgments, young children base orientation judgments almost entirely on visual information.

DISCUSSION

Mammalian sensory systems are not mature at birth, but become increasingly refined as the animal develops. Some basic visual and tactile properties, like contrast sensitivity and acuity reach near-adult levels within the first year of life [5, 6]; while other attributes, like form [7] and motion perception [8, 9] and visual or haptic recognition of 3D objects [10] continue to develop through the school years until 8-14 years of age. The results of this study show that cross-modal integration of form information also develops late: before 8 years of age, children do not integrate visual and haptic spatial information, but one or the other sense dominates, irrespective of its reliability (as assessed by discrimination thresholds), at least over the range we studied. However, there is no evidence that either vision or touch acts as a “gold standard”, always dominating the other. For size discrimination, haptic information dominated in determining not only the perceived height, but also in determining thresholds (a *loser take all* strategy). This would be consistent with ideas going back to Berkeley [11] that “touch educates vision”. But the second experiment did not confirm this trend: for orientation discriminations vision dominated in conditions where vision and haptic information should be weighted roughly equally.

At first sight our results may seem to be at variance with many studies showing that young children and even infants possess a variety of multi-sensory abilities [12]. However, most of these studies do not measure integration *per se*, but the capacity to compare information from different senses. Other studies have demonstrated age-dependent sensory dominance in size-matching, that varies with age up to about twelve, generally with vision dominating young children [eg 13, 14, 15] but not always [16]. However, these experiments also did not study integration by bimodal presentation, but relied on cross-modal matching, a quite different technique. Furthermore, as thresholds were not measured in their particular conditions, it is difficult to know whether the dominance was predicted by MLE or not.

Physiological studies in cat and monkey also point to delayed development of cross-modal integration. In adult animals, many neurons in the deep layers of superior colliculus show strong, super-linear integration of auditory and visual information [17]. However, the integration-enhanced response is not present in young animals, but develops later, after the unimodal visual and auditory properties are completely mature [18, 19]. This has also been demonstrated in a recent psychophysical study [20], showing late development of integration in humans, well after the unimodal orienting response is well established. 8-10 month-old infants showed significant decreases in response times in orientating towards dual-modality compared with single- modality visuo-auditory sources, whereas younger infants showed no dual-modality decrease in latency (above probability-summation predictions). However, although the integration develops late compared with the orienting response, this simple audio-visual integration develops far earlier than the cross-sensory integration of this study, suggesting a clear dissociation. This is interesting, as it shows that children, even infants, do have the capacity to integrate across modalities; whether they integrate or not seems to depend on the task: there is clear evidence for cross-modal integration for a simple orientating response, while for spatial discriminations of size and orientation, integration does not occur. As different modalities, and indeed different tasks within each modality, develop at different rates, it is to be expected that maturation of cross-modal integration should also be task-dependent, only developing after both relevant modalities are mature. It would be interesting to measure neural activity in children doing visuo-haptic form discriminations to see whether the changes in activity noted in lateral occipital and anterior intra-parietal cortices in adults [21] are absent in young children.

Why should cross-sensory integration of spatial information develop so late? One possibility is that sensory systems involved with spatial perception must recalibrate continuously during development, to take into account physical growth, such as lengthening limbs and digits (affecting haptic judgments and average viewing height), inter-ocular separation (affecting stereoscopic depth), and eye-ball length (affecting retinal size). It is possible that for the developing child calibration is more important than optimizing perception by integration: and if sensory information is integrated, one sense cannot be used to calibrate the other. In addition, the rate of physical growth can vary between sensory systems, causing problems for integration.

But why should haptic information dominate size discriminations and visual information orientation discriminations? Orientation is a primary visual quality that can be gleaned directly from the retinal image, without correction for viewing distance or other variables. Indeed, one of the characterizing properties of neurons in primary visual cortex of primates is their selectivity to orientation [4, 22]. However, for haptic discrimination this information is not encoded directly, but needs to be recovered from the pattern of stimulation of sensor array. It therefore seems sensible that the more direct visual information be used for calibration; when in conflict, it will dominate. For the size discrimination, however, the reverse holds true. For vision, size in external world dimensions is not given directly, but needs to be computed from information about not only the retinal extent of stimulation, but also the distance of the object from the eyes, and its slant. For haptic judgments, the information is more direct, coming from the position of the digits (this will of course require long term calibration, but in the short term may be more stable). Therefore for these judgments the more appropriate calibrator is the haptic system, so it should dominate when there is conflict.

So it may be that during development information from different senses is used to calibrate and fine-tune other senses. The direct haptic size information may assist the visual system in calculating size, from estimates of retinal extent and distance estimates. This would be consistent with old [23] and more recent [24] evidence that children below the age of nine have difficulty with size constancy, underestimating the size of distant objects. On the other hand, orientation judgments, basic to vision, may in some way instruct the haptic system to derive them from the spatial patterning of sensory response. On this view, size and orientation should not be dominated by the more precise information, as the MLE model suggests, but by the more direct and robust source of information, even if this source is less precise in a simple discrimination task. And if the various senses are required for cross calibration, they cannot be combined to increase precision.

Figure captions

Figure 1

Illustrations of stimuli, sample psychometric functions. **A&E** Illustration of the experimental setup for size and orientation discrimination (see Sup. Mat. for more details and movie). **B, C, F & G** Sample psychometric functions showing visual (green symbol) and haptic (red symbol) discrimination of size (B&C) and orientation (F&G) discrimination for four representative children: SB age 10.2 (B); DV age 5.5 (C); AR age 8.7 (E); GF age 5.7 (F). The mean of the curves (50% point) estimates the point of subjective equality and the standard deviation the threshold. **D&H.** Average thresholds (geometric average) for haptic (red symbols), visual (green) and visuo-haptic (dark blue) size and orientation discrimination, together with the average MLE predictions (light blue), as a function of age. The predictions were calculated individually for each subject, then averaged. The tick labeled "blur" shows thresholds for visual stimuli blurred by a translucent screen 19 cm from the blocks (see Sup. Mat.). Error bars here and Fig. C are ±1 SEM.

Figure 2

Example psychometric functions of four children, with various degrees of cross-modal conflict. **A & B** size discriminations: SB age 10.2 (A); DV age 5.5 (B); **C&D** orientation discrimination: AR age 8.7 (C); GF age 5.7 (D). The lower colour-coded arrows show the MLE predictions, calculated from threshold measurements (eq. 1 Sup. Mat.). The black dashed horizontal lines show the 50% performance point, intersecting with the curves at their PSE (shown by short vertical bars). The upper colour-coded arrows indicate the size of the haptic standard in the size condition (A&B) and the orientation of visual standard in the orientation condition (C&D). The older children generally follow the adult pattern, while the 5 year-olds were dominated by haptic information for the size task, and visual information for the orientation task.

Figure 3

Summary data showing PSEs for all subjects for all conflict conditions, plotted against the predictions, for size (A) and orientation (B) discriminations (see Sup. Mat. for more details). Different colors refer to different subjects within each age group. The symbol shapes refer to the level of cross-sensory conflict (Δ): squares 3 mm or 4˚; circles –3 mm or –4˚; upright triangles 0; diamonds 2 mm; inverted triangles –2 mm. Closed symbols refer to the no-blur condition for the size judgments, and vertical orientation judgments; open symbols to modest blur (screen at 19 cm) or oblique orientations; cross in symbols to heavy blur (screen at 39 cm).

Figure 4

A. Proportion of variance (R^2) of the PSE data (Fig. 3) explained by three models: haptic dominance (red symbols), visual dominance (green symbols) and MLE prediction (light blue symbols). A value of 1 means that all the variance was explained by the model, 0 that the model performed as well as the mean, and less than 0 that it performed worse than the mean (see eq. 7 of Sup. Mat.). Values less than –1 were clipped for graphical representation (some were as low as –8). The tick labeled "blur" shows the fit to all adult data, unblurred and with the two different levels of blur (see Sup. Mat.) – otherwise the visual stimuli were unblurred. **B.** Haptic and visual weights for the size and orientation discrimination, derived from thresholds via the MLE model (eq. 3 of Sup. Mat.: violet circles) or from PSE values (eq. 6 of Sup. Mat.: black squares). Weights were calculated individually for each subject, then averaged. After 8-10 years the two estimates converge, suggesting that the system is integrating in a statistically optimal manner.

AKNOWLEDGEMENTS

We thank Marc Ernst for valuable advice, and all the children for their willing participation in this research supported by Italian Ministry of Universities and Research, and EC project "RobotCub" (FP6-4270).

REFERENCES

1. Ernst, M.O., and Banks, M.S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. Nature *415*, 429-433.
2. Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. Curr Biol *14*, 257-262.
3. Helbig, H.B., and Ernst, M.O. (2007). Optimal integration of shape information from vision and touch. Exp Brain Res *179*, 595-606.
4. Hubel, D.H., and Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. J. Physiol. (London) *195*, 215-243.
5. Atkinson, J. (2000). The developing visual brain (New York: Oxford University Press).
6. Streri, A. (2003). Cross-modal recognition of shape from hand to eyes in human newborns. Somatosens Mot Res *20*, 13-18.
7. Kovacs, I., Kozma, P., Feher, A., and Benedek, G. (1999). Late maturation of visual spatial integration in humans. Proc Natl Acad Sci U S A *96*, 12204-12209.
8. Atkinson, J., King, J., Braddick, O., Nokes, L., Anker, S., and Braddick, F. (1997). A specific deficit of dorsal stream function in Williams' syndrome. Neuroreport *8*, 1919-1922.
9. Ellemberg, D., Lewis, T.L., Meghji, K.S., Maurer, D., Guillemot, J.P., and Lepore, F. (2003). Comparison of sensitivity to first- and second-order local motion in 5-year-olds and adults. Spat Vis *16*, 419-428.
10. Rentschler, I., Juttner, M., Osman, E., Muller, A., and Caelli, T. (2004). Development of configural 3D object recognition. Behav Brain Res *149*, 107-111.
11. Berkeley, G. (1709). An essay towards a new theory of vision (Indianapolis: 1963: Bobbs-Merril).
12. Lewkowicz, D.J. (2000). Perceptual development: Visual, auditory, and speech perception in infancy. Am J Psychol *113*, 488-500.
13. Misceo, G.F., Hershberger, W.A., and Mancini, R.L. (1999). Haptic estimates of discordant visual-haptic size vary developmentally. Percept Psychophys *61*, 608-614.

14. Hatwell, Y. (1987). Motor and cognitive functions of the hand in infacy and childhood. International journal of behaviour Development *10*, 509-526.

15. McGurk, H.P., Roderick P. (1980). Intermodal coordination in young children: Vision and Touch. Developmental Psychology *16*, 679-680.

16. Klein, R.E. (1966). A developmental study of perception under conditions of conflicting sensory cues. pp. 2162B-2163B.: University Microfilms No. 66-12213.

17. Stein, B.E., Meredith, M.A., and Wallace, M.T. (1993). The visually responsive neuron and beyond: multisensory integration in cat and monkey. Prog Brain Res *95*, 79-90.

18. Stein, B.E., Labos, E., and Kruger, L. (1973). Sequence of changes in properties of neurons of superior colliculus of the kitten during maturation. J Neurophysiol *36*, 667-679.

19. Wallace, M.T., and Stein, B.E. (2001). Sensory and multisensory responses in the newborn monkey superior colliculus. J Neurosci *21*, 8886-8894.

20. Neil, P.A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D.J., and Shimojo, S. (2006). Development of multisensory spatial integration and perception in humans. Dev Sci *9*, 454-464.

21. Helbig, H.B., Ricciardelli, P., Pietrini, P., and Ernst, M.O. (2006). Integration of shape information from vision and touch: Optimal perception and neural correlates. Neuroimage *31*, Sup 1.

22. Tootell, R.B., Hadjikhani, N.K., Vanduffel, W., Liu, A.K., Mendola, J.D., Sereno, M.I., and Dale, A.M. (1998). Functional analysis of primary visual cortex (V1) in humans. Proc Natl Acad Sci U S A *95*, 811-817.

23. Zeigler, H.P., and Leibowitz, H. (1957). Apparent visual size as a function of distance for children and adults. Am J Psychol *70*, 106-109.

24. Granrud, C.E., and Schmechel, T.T. (2006). Development of size constancy in children: a test of the proximal mode sensitivity hypothesis. Percept Psychophys *68*, 1372-1381.

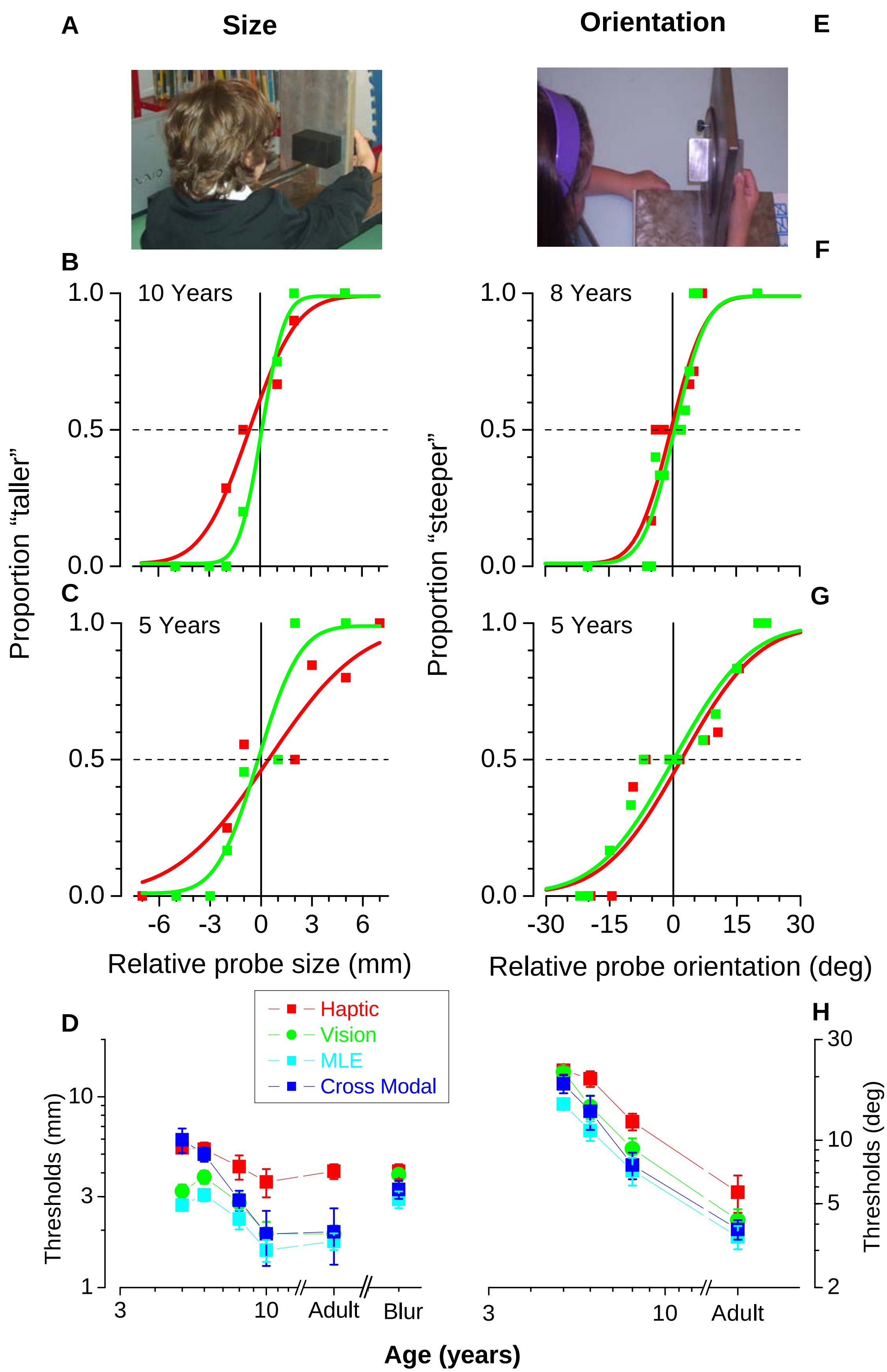
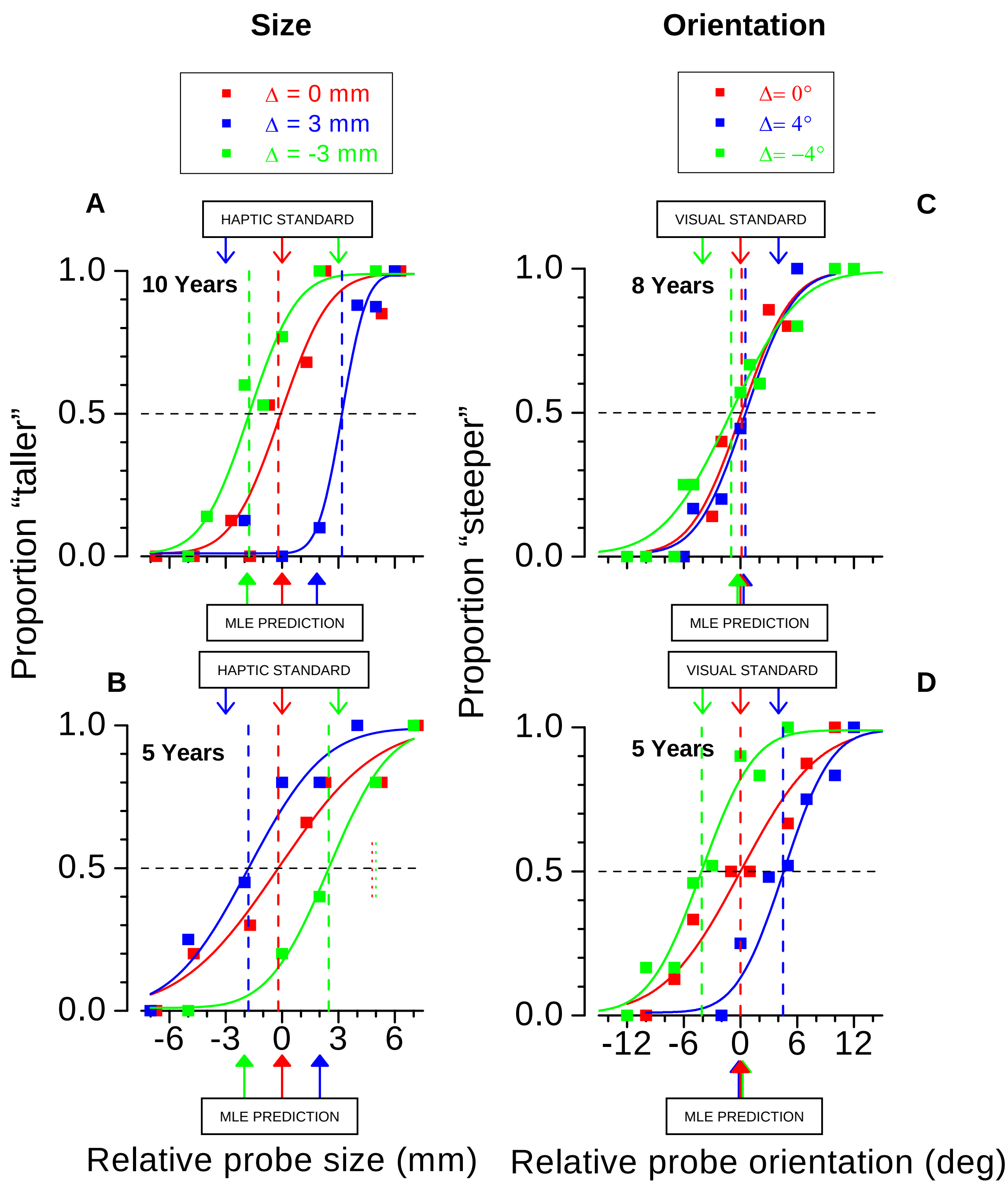
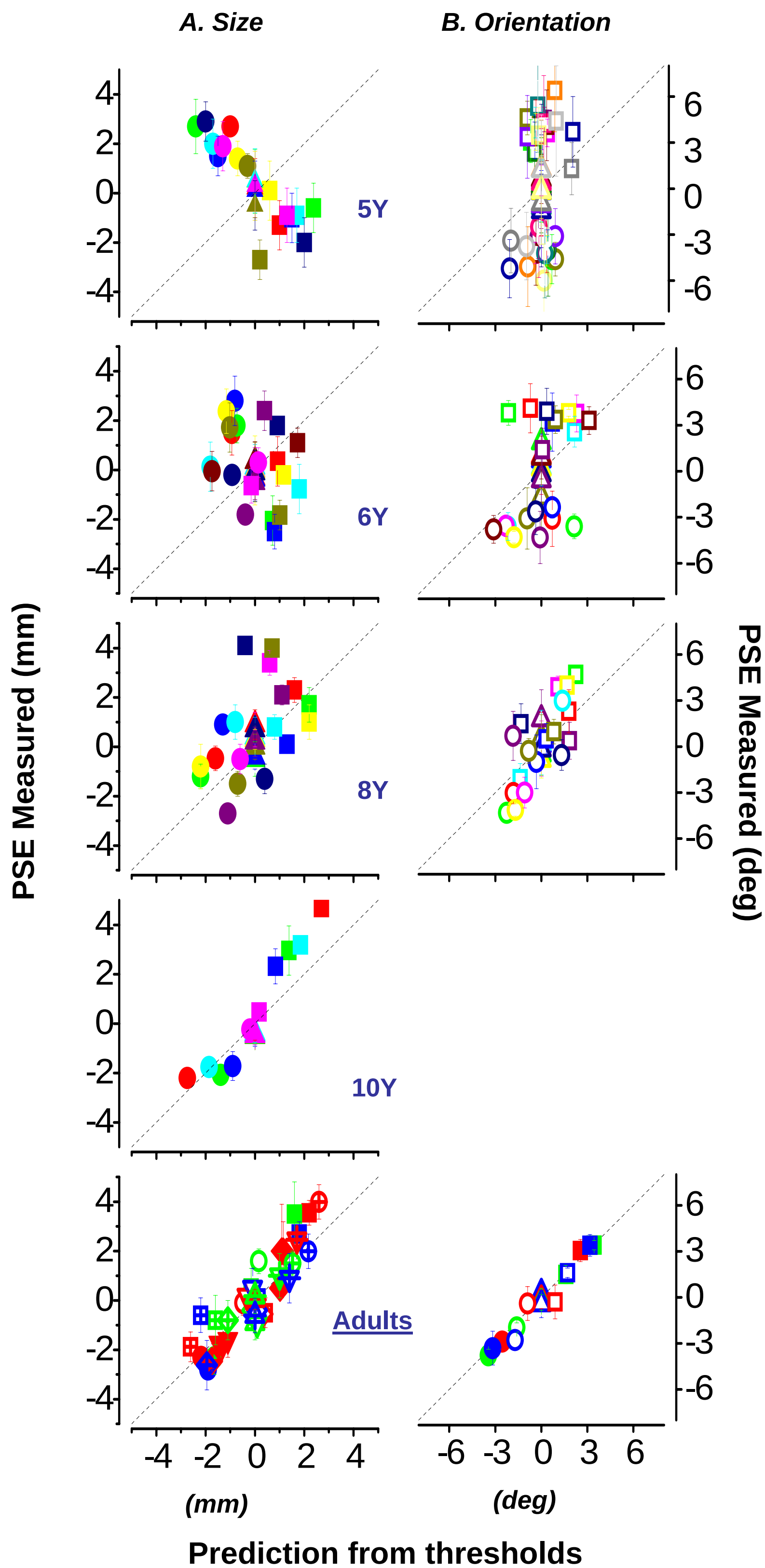

A
Size
Orientation
E
B
10 Years
F
8 Years
1.0
0.5
0.0
Proportion "taller"
Proportion "steeper"
C
5 Years
G
5 Years
-6
-3
0
3
6
-30
-15
0
15
30
Relative probe size (mm)
Relative probe orientation (deg)
D
H
Haptic
Vision
MLE
Cross Modal
Thresholds (mm)
Thresholds (deg)
10
3
1
30
10
5
2
3
10
Adult
Blur
Age (years)

Fig 1



Size
Orientation
Δ = 0 mm
Δ = 3 mm
Δ = -3 mm
Δ= 0°
Δ= 4°
Δ= −4°
A
B
C
D
HAPTIC STANDARD
VISUAL STANDARD
MLE PREDICTION
10 Years
5 Years
8 Years
5 Years
Proportion "taller"
Proportion "steeper"
1.0
0.5
0.0
-6
-3
0
3
6
-12
-6
0
6
12
Relative probe size (mm)
Relative probe orientation (deg)

Fig 2



A. Size
B. Orientation
5Y
6Y
8Y
10Y
Adults
PSE Measured (mm)
PSE Measured (deg)
(mm)
(deg)
Prediction from thresholds

Fig 3

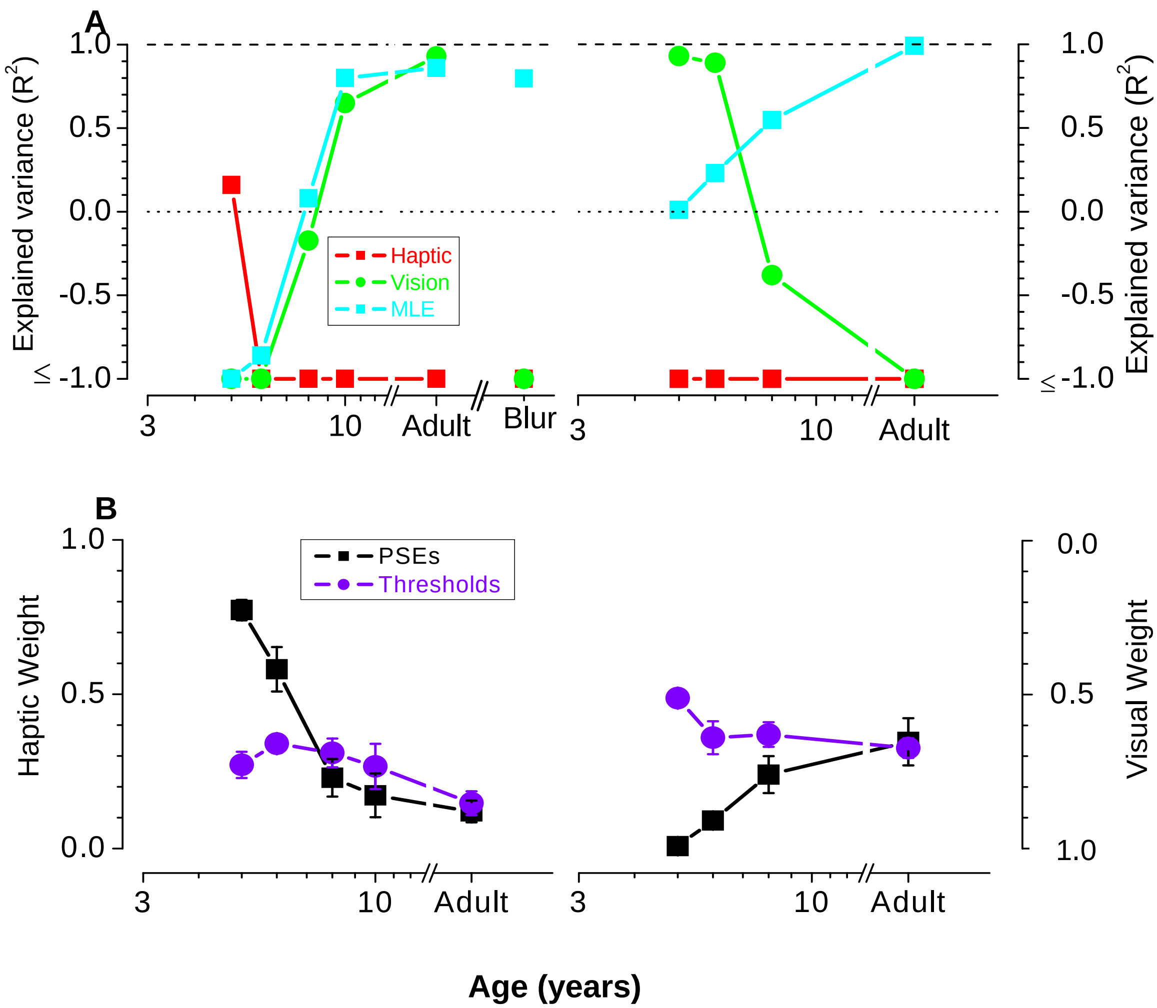

A
B
Explained variance (R²)
1.0
0.5
0.0
-0.5
≤ -1.0
Haptic
Vision
MLE
3
10
Adult
Blur
Haptic Weight
Visual Weight
PSEs
Thresholds
Age (years)

Fig 4