Early binocular input is critical for development of audiovisual

but not visuotactile simultaneity perception

Yi-Chuan Chen[1,2], Terri L. Lewis[1,3], David I. Shore[1], & Daphne Maurer[1,3]

[1]Department of Psychology, Neuroscience & Behaviour, McMaster University, 1280 Main Street West, Hamilton, Ontario, L8S 4K1, Canada

[2]Department of Experimental Psychology, University of Oxford, 9 South Parks Road. Oxford, OX1 3UD, UK

[3]Department of Ophthalmology and Vision Sciences, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, M5G 1X8 Canada

Submitted to: Current Biology (Report)

Word count: 2619 words

Submitted date: 01 Jan 2017

Corresponding author:

Daphne Maurer

Department of Psychology, Neuroscience & Behaviour. McMaster University

1280 Main St West. Hamilton, ON, L8S 4K1, Canada

email: maurer@mcmaster.ca

Summary

Temporal simultaneity provides an essential cue for integrating multisensory signals into a unified perception. Early visual deprivation, in both animals and humans, leads to abnormal neural responses to audiovisual signals in subcortical and cortical areas [1–5]. Behavioural deficits in integrating complex audiovisual stimuli in humans are also observed [6,7]. It remains unclear if early visual deprivation affects visuotactile perception similarly to audiovisual perception, and whether the consequences for either pairing differ after monocular versus binocular deprivation [8–11]. Here, we evaluated the impact of early visual deprivation on the perception of simultaneity for audiovisual and visuotactile stimuli in humans. We tested patients born with dense cataracts in one or both eyes that blocked all patterned visual input until the cataractous lenses were removed and the affected eyes fitted with compensatory contact lenses (mean duration of deprivation = 4.4 months; range = 0.3 to 28.8 months). Both monocularly- and binocularly-deprived patients demonstrated lower precision in judging audiovisual simultaneity. However, qualitatively different outcomes were observed for the two patient groups: the performance of monocularly-deprived patients matched that of young children at immature stages, whereas that of binocularly-deprived patients did not match any stage in typical development. Surprisingly, patients performed normally in judging visuotactile simultaneity after either monocular or binocular deprivation. Therefore, early binocular input is necessary to develop normal neural substrates for simultaneity perception of visual and auditory events, but not visual and tactile events.

Keywords: Multisensory perception, Audiovisual simultaneity, Visuotactile simultaneity, Visual deprivation, Early visual experience, Congenital cataract

Results and Discussion

We examined how a short period of monocular or binocular deprivation after birth affects the later perception of simultaneity for audiovisual and visuotactile stimuli in humans. To do so, we compared the simultaneity judgements of patients treated for congenital cataracts during infancy in one or both eyes to those of age-matched visually normal controls. Participants reported whether a flash and a beep (or a tap) were perceived simultaneously when presented at 15 different stimulus onset asynchronies (SOAs) [12]. All patients were at least 11 years of age at testing (i.e., they had at least 10 years of post-treatment visual input), which is after the age of normal maturation for the perception of both audiovisual and visuotactile simultaneity (at 9 and 11 years of age, respectively [13,14]).

For audiovisual events, all patients judged the two stimuli to be simultaneous over a wider range of SOAs than did controls (indexed by δ, half of the width of the response distribution at half height, see Figure 1 and Table 1; see Supplemental data analysis), but the pattern of deviation differed for binocularly- and monocularly-deprived patients. For binocularly-deprived patients, the differences were only on trials when the flash came before the sound (Figure 1A). This was reflected statistically in more errors in the visual-leading condition (εVF), and a larger shift of the point of subjective simultaneity (the midpoint of the distribution) toward the visual-leading side than in controls. The small shift toward the visual-leading side in typical adults (Figure 2B) can be explained, at least in part, by the fact that processing time is shorter for auditory than visual stimuli, but the shift was larger for binocularly-deprived patients than controls (indexed by the more negative τ). This pattern is different from any stage in the typical developmental trajectory [13], and hence reflects an atypical development.

In contrast, the monocularly-deprived patients demonstrated a wider audiovisual simultaneity window than controls in both visual-leading and auditory-leading conditions, irrespective of the eye tested (Figures 1B & 1C). This was reflected in a higher threshold of audiovisual simultaneity (δ), with no significant shift of the point of subjective simultaneity than in controls (Figure 2). They were also *less* likely than controls to make response errors on the 0-ms trials (εS), at least when using the deprived eye. Interestingly, there were no significant differences between the deprived and fellow eye on any measure (Table 1, all $t(12) < 1.68$, ps > .12). Thus, early monocular deprivation produces a wider-than-normal window of audiovisual simultaneity for both the deprived and the fellow eye. This pattern is similar to that of children with normal eyes who have not yet reached adult levels of precision [13].

The results were strikingly different for the visuotactile events. Both groups of patients were entirely normal, despite viewing the same visual stimuli as in the audiovisual task. Neither binocularly-deprived (Figure 1D) nor monocularly-deprived (Figures 1E & 1F) patients had a wider visuotactile simultaneity window or a shift in the point of subjective simultaneity compared to the controls. The only difference was *fewer* response errors in the 0-ms trials (εS) in the monocularly-deprived patients than controls (i.e., the peak of the distribution was higher) when using either the deprived or the fellow eye (Table 2). The patients' normal perception of visuotactile simultaneity is surprising, since evidence from an animal study demonstrates that early deprivation in a given sensory modality prevents the normal development of multisensory neurons associated with that modality [15].

Combined, the results from binocularly- and monocularly-deprived patients indicate that early visual input to *both* eyes is essential to set up the neural architecture for the development of a normal *audiovisual* simultaneity window, consistent with the results of animal studies [16,17]. If visual input is to only one eye, it can partially tune the deprived eye, resulting in a typical but immature outcome for both the deprived eye *and* the fellow eye. However, the amount of lifetime perceptual experience after infancy (at least beyond 11 years of age—the youngest age tested) did not determine performance (see Table S3), contrary to the evidence from multisensory neurons in adult cats' superior colliculus [2].

The larger time difference between auditory and visual processing for binocularly-deprived patients than controls (the more negative τ) is likely caused by faster-than-normal auditory processing rather than slower visual processing, given that patients' *visuotactile* simultaneity perception and temporal vision are spared [10,18]. The fact that the abnormality manifested on only the visual-leading side of the audiovisual simultaneity window can be explained by two non-mutually exclusive mechanisms. The first builds on the finding that attention of binocularly-deprived patients is biased toward audition over vision [19], unlike that of typically-developing adults who preferentially attend vision [20]. Signals in an attended modality are known to be processed faster (i.e., the prior entry effect [21,22]). Hence, the binocularly-deprived patients' perception of audiovisual simultaneity would be most likely to deviate from normal when the visual flash was presented earlier than the beep, as was found here.

The second account is based on adaptation to the natural condition that light travels faster than sound, so that the visual signal typically arrives earlier than the auditory signal when originating from the same distal event. Hence, the visual-leading side is the condition to which people adapt in daily life [23], and it is the condition susceptible to perceptual training in visually normal adults [24,25]. Early visual deprivation may prevent such natural adaptation from developing optimally. This conjecture is based on the evidence that in binocularly-deprived patients, crossmodal adaptation for motion perception is influenced more-than-normal by auditory compared to visual stimuli [26]. The abnormalities in both the audiovisual attention and adaptation systems in binocularly-deprived patients may arise from unbalanced Hebbian competition between visual and auditory inputs before treatment, leading to the preservation, or recruitment, of auditory signaling in cortical areas that are typically visually dominant [4,5], as occurs in the congenitally blind [27,28]. However, normal visual input to the fellow eye of monocularly-deprived patients before treatment may allow a normal setting-up of the audiovisual system, given that neurons in superior temporal sulcus, an area central to audiovisual simultaneity [29-32] receive visual information from each eye. That early input from one eye may be sufficient to set up a normal substrate for audiovisual simultaneity, albeit not one that becomes as precisely tuned as is achieved after normal early input to both eyes.

The audiovisual abnormalities cannot simply be the consequence of abnormal visual development since both binocularly- and monocularly-deprived patients have normal simultaneity perception of the same visual flash when paired with a tactile tap. That rudimentary forms of audiovisual interactions are observed at birth [33,34] but visuotactile interactions develop postnatally with onset at about 6-9 months of age [35–37] leads to the expectation from previous unimodal visual studies that the neural substrates for visuotactile interactions should have been more likely to be impaired by early visual deprivation (i.e., the sleeper effect [10,38]). However, the reversed developmental outcomes indicate that the roles of early visual experience for the development of the visual system and for multisensory systems differ. The mechanism of crossmodal calibration during development provides a more likely alternative explanation [39,40].

The constant calibration hypothesis suggests that the modality with more accurate perception (though not necessarily more precision) calibrates the other during the development of multisensory systems [40,41]. It leads to the conclusion that the visuotactile system is calibrated by touch, whereas the audiovisual system, by vision. For the visuotactile system, the analysis is based on the observation that most tactile events are presented on the skin and that perceiving tactile events often stems from reaching and grasping objects in peripersonal space. The location and timing of these events are accurately encoded and updated with real-time feedback by the somatotopic and proprioceptive systems. This analysis is consistent with evidence that touch calibrates vision during their interactions [42]. In contrast, audiovisual events are often distal: visual signals arrive in a virtual instant regardless of distance, whereas for auditory signals arrival takes time, and more time is required as distance is increased. In typically developing adults, adaptation to audiovisual asynchrony causes a shift in perceived timing, such that the auditory signals are shifted toward the visual signal [43]. Further, the refinement of the audiovisual simultaneity window depends on calibration in terms of the arrival time difference between visual and auditory signals as a function of stimulus distance [44,45]. The patients' relatively poor vision, and particularly their poor stereo vision that is important for estimating stimulus distance (Tables S1 & S2), may prevent fine tuning of the audiovisual simultaneity window, whereas their normal tactile perception allows normal fine tuning of visuotactile simultaneity.

Even though we cannot be certain that our results reflect abnormal audiovisual *integration*, rather than merely temporal matching between the visual and auditory signal [13], our results are consistent with previously reported abnormalities of audiovisual perception in binocularly-deprived patients. For example, Putzar et al. [6] reported that, for binocularly-deprived patients, compared to controls, a task-irrelevant tone was less likely to temporally capture a visual target when the visual target was leading. This result can be explained by the binocularly-deprived patients' lower precision of audiovisual simultaneity in the visual-leading side demonstrated in the present study. In addition, Putzar et al. [6,7] demonstrated that binocularly-deprived patients were less likely than controls to integrate visual and auditory speech information. This is consistent with the correlation in typical adults between a *wider* audiovisual simultaneity window in the visual-leading side and *smaller* magnitude of audiovisual speech integration [46,47]. It is possible that, because binocularly-deprived patients are less sensitive than controls to audiovisual simultaneity, it is harder for them to correctly bind visual and auditory signals that should go together, especially when the visual signal comes first.

In summary, we demonstrated that early visual input to both eyes is critical for setting up the neural architecture for the later development of audiovisual but not visuotactile simultaneity. Thus, normal early visual experience is necessary for some but not all types of multisensory integration associated with vision.

Experimental Procedures

Participants

Visually deprived participants were 14 patients treated for binocular congenital cataract (mean age = 23.1 years; age range = 14–34 years), and 15 patients treated for monocular congenital cataract (mean age = 21.8 years; age range = 11–43 years). Eleven binocularly-deprived patients and ten monocularly-deprived patients took part in both audiovisual and visuotactile experiments (see Tables S1 & S2 for details; see Supplemental patients' history of visual deprivation and treatments).

For each visually deprived patient, we tested three age- and gender-matched visually normal controls in the audiovisual simultaneity judgments task, and two controls in the visuotactile simultaneity judgments task. For patients younger than 18 years old, their age-matched controls were in the range of ±1 year, while for those 18 years or older, the range was ±3 years. All controls had no history of eye problems other than refractive errors and all met our criteria on a visual-screening exam (20/20 corrected vision on the Lighthouse eye chart, a stereoacuity of at least 50 arcsec on the Titmus test, and normal binocular fusion in the Worth 4 Dot test). The dominant eye of the controls was determined using the Miles test.

We tested the eye with better acuity in binocularly-deprived patients and compared it to the dominant eye of their paired controls. We tested each eye separately of monocularly-deprived patients, comparing patients' deprived eye to the non-dominant eye of controls, and patients' fellow eye to the dominant eye of controls. For monocularly-deprived patients, we counterbalanced the order of eye tested first and followed the same order for their matched controls.

Adult participants and parents of children provided written consent; children provided written assent. All of the participants were naïve regarding the purpose of the study. The study was approved by the Research Ethics Boards of McMaster University and The Hospital for Sick Children, and conformed to the Tri-Council Statement on Ethical Conduct of Research Involving Humans (TCPS2) (Canada).

Apparatus and Stimuli

Participants were seated in a dimly-lit room with their head on a chin rest located 50 cm from the monitor (24-inch LED monitor with 1920 x 1080 resolution) where the visual stimuli were presented. A gray ring with a 2° inner diameter and 0.6° thickness was displayed in the center of a black background throughout the experiment. The visual stimulus was a 2° white disc presented in the center of the grey ring for 17 ms (1 frame at the 60 Hz refresh rate). The auditory stimulus consisted of a beep presented from speakers placed on either side of the monitor. A beep consisted of white noise with a loudness of 57.5 dB SPL (the room had a background noise level of 40 dB SPL). The duration of the beep was 17 ms with 2 ms on and off ramping. The tactile stimulus was a tap induced by a pin moving up and down (lasting 17 ms) generated by a mechanical device. Participants were asked to place the index finger of their dominant hand on top of the tactile device which was placed in front of the monitor and below the ring so as to align participant's body midline. Presentation of the stimuli was controlled by Matlab (MathWorks Inc., Natick, MA) and Psychtoolbox extensions [48-50].

Design and procedure

In the audiovisual simultaneity judgments task, the method of constant stimuli was used to present the auditory beep and visual flash at each of 15 SOAs (±500, ±400, ±300, ±200, ±150, ±100, ±50, and 0 ms), where negative values indicate that the auditory beep was presented first, and positive values indicate that the visual flash was presented first. The SOA between the flash and beep was confirmed with an oscilloscope. Each SOA was tested twice in each block and participants completed 10 blocks for each tested eye. Participants were allowed to take a short break between blocks and typically took 30 minutes to complete the experiment for one eye.

During the audiovisual experiment, the participants were instructed to fixate the ring. The task was to press "1" if they considered that the flash and beep were presented at the same time or "2" if they considered that they were presented at different times. For participants younger than 15 years of age, responses were reported orally and an experimenter sitting beside the participant keyed the answers into the computer.

In the visuotactile simultaneity judgments task, the SOAs and other design features were the same except that the beep was replaced by the tap. During the experiment, participants heard a continuous white noise presented from a closed-ear headphone in order to mask the noise produced by the tactile device. All of the participants reported their judgment orally and an experimenter keyed answers into the computer. The audiovisual and visuotactile simultaneity tasks were completed in separate sessions, and those who participated in both tasks (21 patients, see above) completed the audiovisual task first.

Two practice sessions for each of the audiovisual and visuotactile tasks were conducted prior to the main experiment. For each task, the first practice session consisted of eight trials: four with 0 ms SOA and one for each of ±500 or ±300 ms SOA. All of the participants needed to achieve 85% accuracy (i.e., no more than one error) in order to proceed and were given three chances to meet this criterion. More than 90% of the participants met the criterion on the first run, and the remaining participants met the criterion on the second run. The second practice session consisted of one trial for each of the 15 SOAs used in the main experiment and was designed to familiarize participants with the experimental procedure. There were no accuracy requirements and no feedback for the second practice session.

References

1. Carriere, B. N., Royal, D. W., Perrault, T. J., Morrison, S. P., Vaughan, J. W., Stein, B. E., and Wallace, M. T. J. (2007). Visual deprivation alters the development of cortical multisensory integration. Neurophysiol. *98,* 2858–2867.
2. Yu, L., Rowland, B. A., and Stein, B. E. (2010). Initiating the development of multisensory integration by manipulating sensory experience. J. Neurosci. *30,* 4904–4913.
3. Stein, B. E., Stanford, T. R., and Rowland, B. A. (2014). Development of multisensory integration from the perspective of the individual neuron. Nat. Rev. Neurosci. *15,* 520–535.
4. Collignon, O., Dormal, G., de Heering, A., Lepore, F., Lewis, T. L., and Maurer, D. (2015). A short postnatal period of visual deprivation triggers long-lasting crossmodal reorganization of the occipital cortex in humans. Curr. Biol. *25,* 2379–2383.
5. Guerreiro, M. J., Putzar, L., and Röder, B. (2015). The effect of early visual deprivation on the neural bases of multisensory processing. Brain *138,* 1499–1504.
6. Putzar, L., Goerendt, I., Lange, K., Rösler, F., and Röder, B. (2007). Early visual deprivation impairs multisensory interactions in humans. Nat. Neurosci. *10,* 1243–1245.
7. Putzar, L., Hötting, K., and Röder, B. (2010). Early visual deprivation affects the development of face recognition and of audio-visual speech perception. Restor. Neurol. Neuros. *28,* 251–257.
8. Ellemberg, D., Lewis, T. L., Maurer, D., and Brent, H. P. (2000). Influence of monocular deprivation during infancy on the later development of spatial and temporal vision. Vision Res. *40,* 3283–3295.
9. Ellemberg, D., Lewis, T. L., Maurer, D., Brar, S., and Brent, H. P. (2002). Better perception of global motion after monocular than after binocular deprivation. Vision Res. *42,* 169–179.
10. Maurer, D., Lewis, T. L., and Mondloch, C. J. (2005). Missing sights: Consequences for visual cognitive development. Trends. Cogn. Sci. *9,* 144–151.

11. Lewis, T. L., and Maurer, D. (2009). Effects of early pattern deprivation on visual development. Optometry Vision Sci. *86*, 640–646.

12. Zampini, M., Guest, S., Shore, D. I., and Spence, C. (2005). Audio-visual simultaneity judgments. Percept. Psychophys. *67*, 531–544.

13. Chen, Y.-C., Shore, D. I., Lewis, T. L., and Maurer, D. (2016). The development of the perception of audiovisual simultaneity. J. Exp. Child Psychol. *146,* 17–33.

14. Chen, Y.-C., Lewis, T. L., Shore, D. I., and Maurer, D. (2015). The development of the perception of visuotactile simultaneity. Perception *44(Sl),* 341–342.

15. Xu, J., Yu, L., Stanford, T. R., Rowland, B. A., and Stein, B. E. (2015). What does a neuron learn from multisensory experience? J. Neurophysiol. *113,* 883–889.

16. Wallace, M. T., Perrault, T. J., Hairston, W. D., and Stein, B. E. (2004). Visual experience is necessary for the development of multisensory integration. J. Neurosci. *24*, 9580–9584.

17. Xu, J., Yu, L., Rowland, B. A., Stanford, T. R., and Stein, B. E. (2012). Incorporating cross-modal statistics in the development and maintenance of multisensory integration. J. Neurosci. *32,* 2287–2298.

18. Ellemberg, D., Lewis, T. L., Maurer, D., Lui, C. H., and Brent, H. P. (1999). Spatial and temporal vision in patients treated for bilateral congenital cataracts. Vision Res. *39,* 3480–3489.

19. de Heering, A., Dormal, G., Pelland, M., Lewis, T. L., Maurer, D., and Collignon, O. (2016). A brief period of postnatal visual deprivation alters the balance between auditory and visual attention. Curr. Biol. *26,* 3101–3105.

20. Posner, M. I., Nissen, M. J., and Klein, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. Psychol. Rev. 83, 157–171.

21. Spence, C., Shore, D. I., and Klein, R. M. (2001). Multisensory prior entry. J. Exp. Psychol. Gen. *130,* 799–832.

22. Zampini, M., Shore, D. I., and Spence, C. (2005). Audiovisual prior entry. Neurosci. Lett. *381*, 217–222.

23. Vroomen, J., and Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. Atten. Percept. Psycho. *72,* 871–884.

24. Powers, A. R., Hillock, A. R., and Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding. J. Neurosci*.* *29,* 12265–12274.

25. Donohue, S. E., Woldorff, M. G., and Mitroff, S. R. (2010). Video game players show more precise multisensory temporal processing abilities. Atten. Percept. Psycho. *72,* 1120–1129.

26. Guerreiro, M. J., Putzar, L., and Röder, B. (2016). Persisting cross-modal changes in sight-recovery individuals modulate visual perception. Curr. Biol. *26,* 3096–3100.

27. Bavelier, D., and Neville, H. J. (2002). Cross-modal plasticity: Where and how? Nat. Rev. Neurosci. *3,* 443–452.

28. Pascual-Leone, A., Amedi, A., Fregni, F., and Merabet, L. B. (2005). The plastic human brain cortex. Annu. Rev. Neurosci. *28,* 377–401.

29. Calvert, G. A., Campbell, R., and Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Curr. Biol. *10,* 649–657.

30. Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyn, J. H., and Martin, A. (2004). Unraveling multisensory integration: Patchy organization within human STS multisensory cortex. Nat. Neurosci. *7,* 1190–1192.

31. Noesselt, T., Rieger, J. W., Schoenfeld, M. A., Kanowski, M., Hinrichs, H., Heinze, H. J., and Driver, J. (2007). Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. J. Neurosci*.* *27*, 11431–11441.

32. Noesselt, T., Bergmann, D., Heinze, H.-J., Münte, T., and Spence, C. (2012). Coding of multisensory temporal patterns in human superior temporal sulcus. Front. Integr. Neurosci. *6*:*64*.

33. Lewkowicz, D. J., Leo, I., and Simion, F. (2010). Intersensory perception at birth: Newborns match nonhuman primate faces and voices. Infancy *15,* 46–60.

34. Morrongiello, B. A., Fenwick, K. D., and Chance, G. (1998). Crossmodal learning in newborn infants: Inferences about properties of auditory-visual events. Infant Behav. Dev. *21,* 543–553.

35. Rose, S. A., Gottfried, A. W., and Bridger, W. H. (1981). Cross-modal transfer in 6-month-old infants. Dev. Psychol. *17,* 661–669.

36. Karl, J. M., and Whishaw, I. Q. (2014). Haptic Grasping configurations in early infancy reveal different developmental profiles for visual guidance of the Reach versus the Grasp. Exp. Brain Res. *232,* 3301–3316.

37. Ali, J. B., Spence, C., and Bremner, A. J. (2015). Human infants' ability to perceive touch in external space develops postnatally. Curr. Biol. *25,* R978–R979.

38. Maurer, D., Mondloch, C. J., and Lewis, T. L. (2007). Sleeper effects. Developmental Sci. *10,* 40–47.

39. Ernst, M. O. (2008). Multisensory integration: A late bloomer. Curr. Biol. *18,* R519–R521.

40. Burr, D., and Gori, M. (2012). Multisensory integration develops late in humans. In The Neural Bases of Multisensory Processes, M. M. Murray & M. T. Wallace, Eds. (Boca Raton, FL: CRC Press), pp. 345–362.

41. Gori, M. (2015). Multisensory integration and calibration in children and adults with and without sensory and motor disabilities. Multisens. Res. *28,* 71–99.

42. Gori, M., Sciutti, A., Burr, D., and Sandini, G. (2011). Direct and indirect haptic calibration of visual size judgments. PLoS ONE *6(10),* e25599.

43. Navarra, J., Hartcher-O'Brien, J., Piazza, E., and Spence, C. (2009). Adaptation to audiovisual asynchrony modulates the speeded detection of sound. P. Natl. Acad. Sci. U.S.A, *106,* 9169–9173.

44. Sugita, Y., and Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. Nature *421,* 911.

45. King, A. J. (2005). Multisensory integration: Strategies for synchronization. Curr. Biol. *15*, R339–R341.

46. Stevenson, R. A., Zemtsov, R. K., and Wallace, M. T. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusion. J. Exp. Psychol. Human *38,* 1517–1529.

47. Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., and Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. J. Neurosci. *34,* 691–697.

48. Brainard, D. H. (1997). The Psychophysics Toolbox. Spatial Vision *10,* 433–436.

49. Pelli, D. G. (1997) .The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spatial Vision *10*, 437–442.

50. Kleiner, M., Brainard, D., and Pelli, D. (2007). What's new in Psychtoolbox-3*?* Perception *36(Sl),* 14.

Figure Legends

Figure 1. The mean percentage of audiovisual (top row) and visuotactile (bottom row) simultaneous responses. Panels (A) and (D) are for binocularly-deprived patients using the better eye and their controls using the dominant eye; panels (B) and (E) are for monocularly-deprived patients using the deprived eye and their controls using the non-dominant eye; panels (C) and (F) are for monocularly-deprived patients using the fellow eye and their controls using the dominant eye. The error bars represent ±1 standard error of the mean. See Tables S1 and S2 for the patients' clinical history.

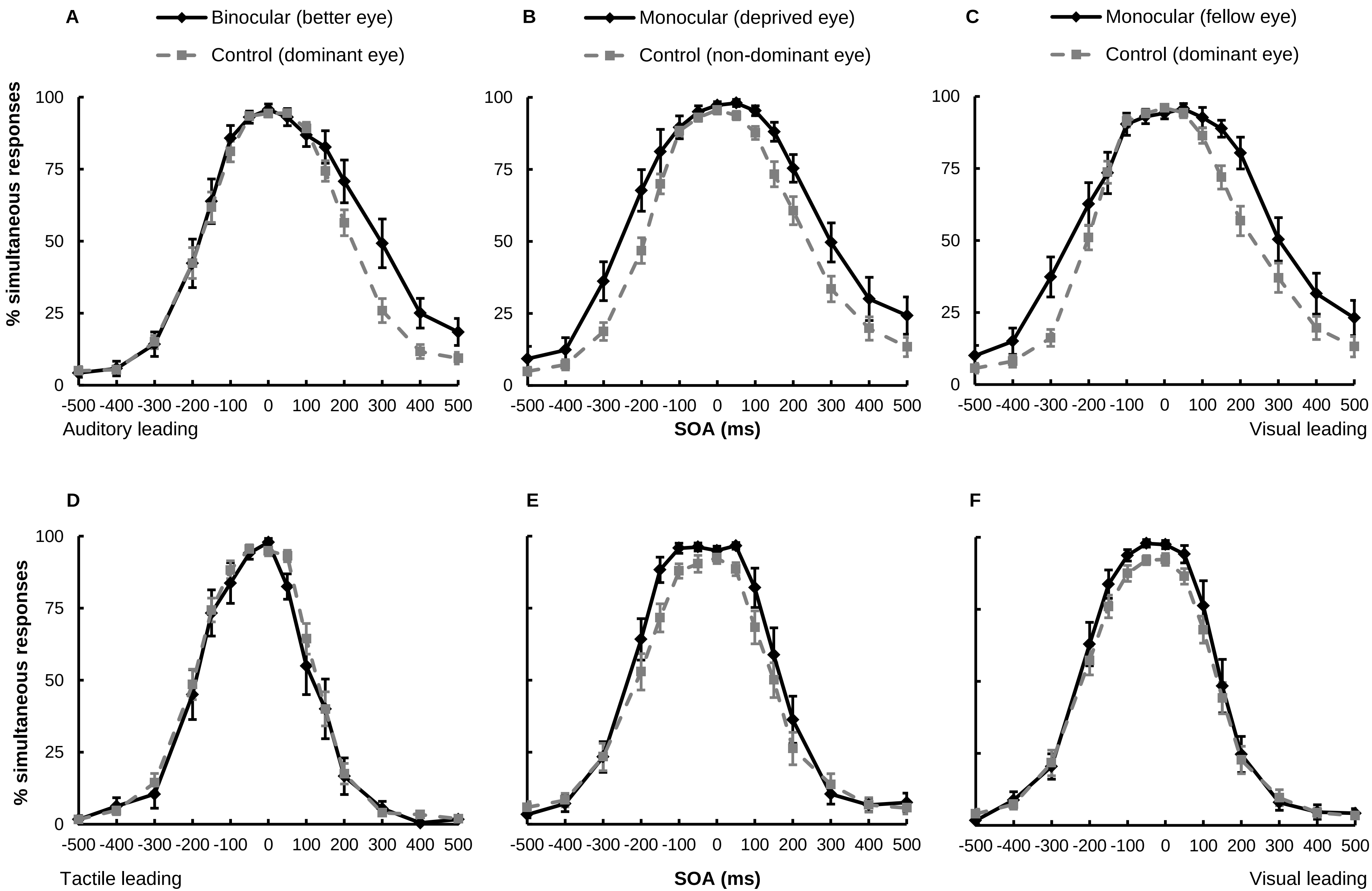
Figure 2. Audiovisual simultaneity judgments. (A) The mean threshold of audiovisual simultaneity (δ) and (B) the mean point of subjective simultaneity for the binocularly-deprived patients (using the better eye) and monocularly-deprived patients (using the deprived and the fellow eyes, respectively), and their paired controls. See Figure S1 for individual patient's results. In the control groups, the point of subjective simultaneity was significantly larger than 0 in all three conditions (all $t(38) > 1.78$, ps $< .05$, one-tailed), suggesting that the point of subjective simultaneity is located on the visual-leading side, as demonstrated in previous studies [12, 13, 23,46].

Author contributions

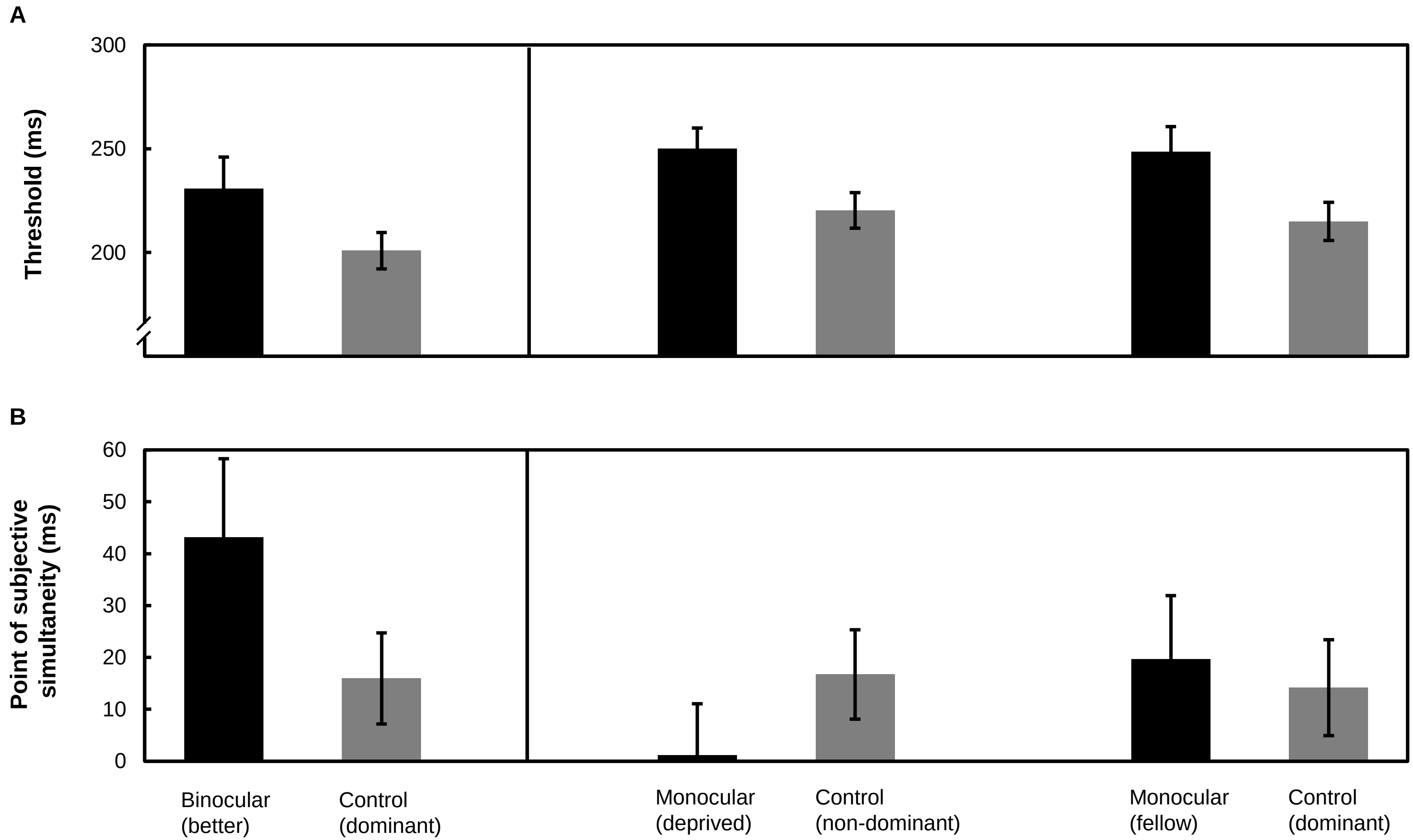
All of the authors designed the study and wrote the paper; YCC conducted the experiments and data analysis.

Acknowledgement

This research was supported by a collaborative activity award from the James S. McDonnell Foundation to DM and a grant from the Natural Sciences and Engineering Council (Canada) [9797] to DM. We thank Sally Stafford for her help with all aspects of the research, and Brendan Stanley, Joseph Capozza, Mina Morcos, Natalia Fong, and Sarah Shaikh for helping with the testing of age-matched controls.



A
Binocular (better eye)
Control (dominant eye)
B
Monocular (deprived eye)
Control (non-dominant eye)
C
Monocular (fellow eye)
Control (dominant eye)
% simultaneous responses
100
75
50
25
0
-500 -400 -300 -200 -100 0 100 200 300 400 500
Auditory leading
SOA (ms)
Visual leading
D
E
F
Tactile leading



A
Threshold (ms)
300
250
200
B
Point of subjective simultaneity (ms)
60
50
40
30
20
10
0
Binocular (better)
Control (dominant)
Monocular (deprived)
Control (non-dominant)
Monocular (fellow)
Control (dominant)

Table 1. The estimated parameters for audiovisual simultaneity judgments for patients. Each parameter (see Supplemental data analysis) was converted into z-score based on the mean and standard deviation of the age- and gender-matched control group. We compared the z-scores of the simultaneity thresholds using one-tailed one-sample t-tests because the patients demonstrated a wider distribution than controls. For all other comparisons, we used two-tailed t-tests. Further comparisons between the better eye of binocularly-deprived patients and the deprived eye of monocularly-deprived patients demonstrated significant differences in the point of subjective simultaneity that was shifted further toward the visual-leading side in binocularly-deprived patients ($t(24) = 2.77$, $p < .05$). This larger shift was attributable partly to the fact that typically the time is shorter for auditory than visual processing, and this difference was larger in binocularly- than monocularly-deprived patients (indexed by τ, $t(24) = 2.21$, $p < .05$). Note that these results cannot be explained by the patients' visual acuity at the time of the test given that visual acuity was superior for the better eye of binocularly-deprived patients (mean LogMar = 0.28) than for the deprived eye of monocularly-deprived patients (mean LogMar = 0.71, $t(24) = 4.25$, $p < .001$). See also Figure S1 and Table S3.

Patient	Binocularly-deprived		Monocularly-deprived			
Tested eye	Better eye		Deprived eye		Fellow eye	
Parameter	**Mean** (SE)	*P*	**Mean** (SE)	*p*	**Mean** (SE)	*p*
Threshold of audiovisual simultaneity (δ)	**0.55** (0.28)	< .05	**0.56** (0.19)	< .05	**0.58** (0.21)	<.05
Point of subjective simultaneity (PSS)	**0.72** (0.24)	< .05	**-0.35** (0.30)	.26	**0.11** (0.32)	.73
Auditory processing variability (λA)	**-0.22** (0.25)	.39	**-0.06** (0.30)	.85	**0.08** (0.25)	.75
Visual processing variability (λV)	**0.29** (0.32)	.38	**-0.24** (0.21)	.28	**-0.31** (0.22)	.18
AV arrival time difference (τ)	**-0.67** (0.23)	< .05	**0.26** (0.35)	.48	**0.06** (0.36)	.86
Response errors of A first (εAF)	**-0.22** (0.12)	.10	**0.33** (0.45)	.49	**0.55** (0.34)	.14
Response errors of simultaneous (εS)	**0.32** (0.43)	.46	**-0.39** (0.12)	< .05	**-0.11** (0.34)	.75
Response errors of V first (εVF)	**0.82** (0.35)	< .05	**0.55** (0.35)	.15	**0.50** (0.30)	.12

A: audition; V: vision

Table 2. The estimated parameters for visuotactile simultaneity judgments for patients. As in the audiovisual task, we compared the z-scores of the simultaneity thresholds using one-tailed t-tests; for all other comparisons, we used two-tailed t-tests.

Patient	Binocularly-deprived		Monocularly-deprived			
Tested eye	Better eye		Deprived eye		Fellow eye	
Parameter	**Mean** (SE)	*P*	**Mean** (SE)	*p*	**Mean** (SE)	*p*
Threshold of visuotactile simultaneity (δ)	**-0.28** (0.39)	= .50	**0.26** (0.24)	= .31	**0.10** (0.23)	= .69
Point of subjective simultaneity (PSS)	**0.22** (0.31)	= .52	**0.16** (0.13)	= .27	**-0.03** (0.23)	= .92
Tactile processing variability (λT)	**0.06** (0.28)	= .85	**-0.14** (0.31)	= .68	**0.78** (0.44)	= .11
Visual processing variability (λV)	**0.69** (0.48)	= .19	**0.70** (0.48)	= .19	**-0.005** (0.23)	= .98
TV arrival time difference (τ)	**-0.25** (0.30)	= .43	**-0.34** (0.18)	= .09	**0.28** (0.22)	= .24
Response errors of T first (εTF)	**0.42** (0.62)	= .52	**-0.21** (0.12)	= .13	**0.25** (0.35)	= .50
Response errors of simultaneous (εS)	**-0.15** (0.24)	= .57	**-0.32** (0.08)	< .005	**-0.56** (0.14)	< .005
Response errors of V first (εVF)	**-0.13** (0.22)	= .59	**0.09** (0.22)	= .69	**0.03** (0.28)	= .93

T: touch; V: vision

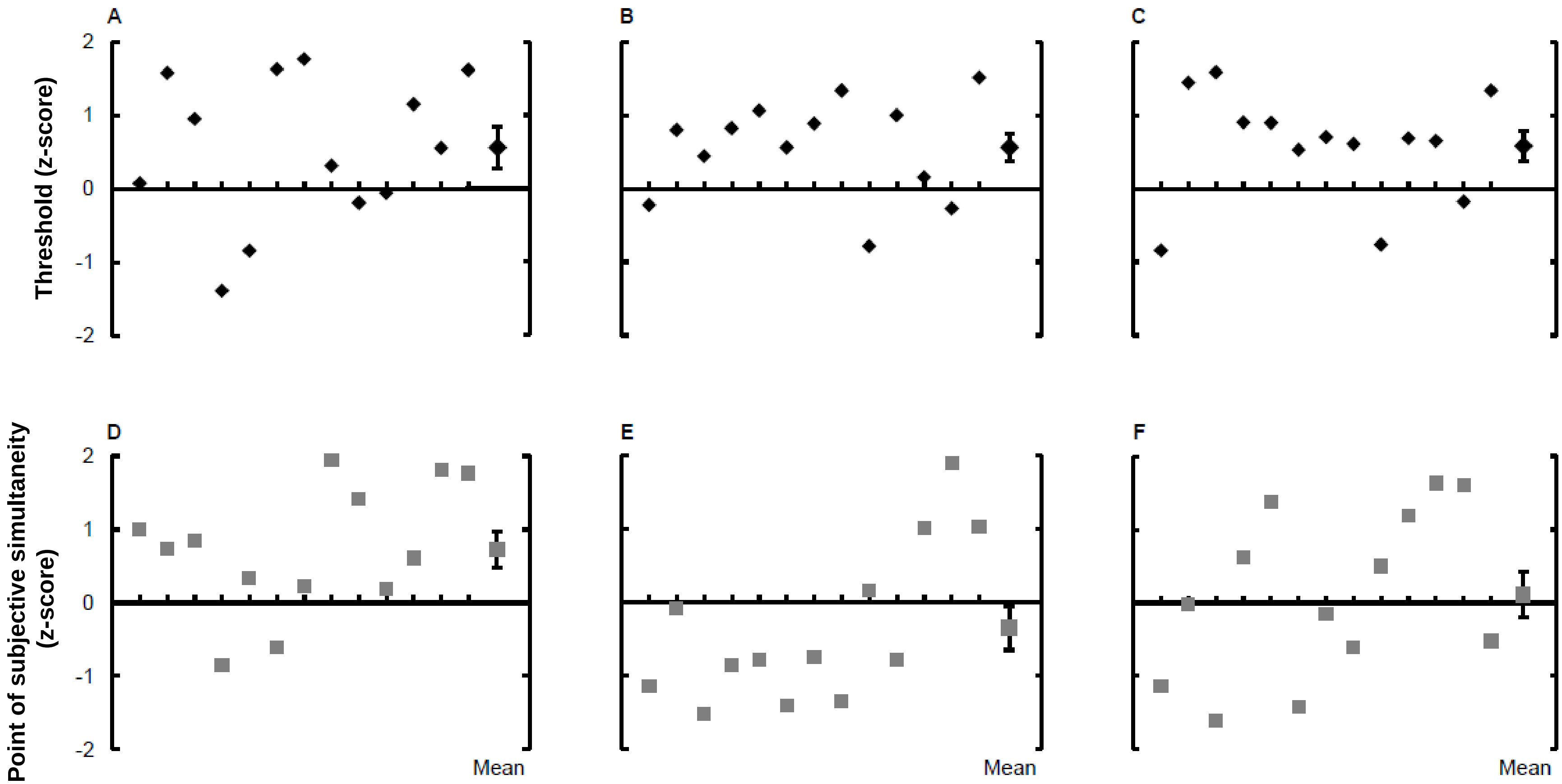


Figure S1. The individual and mean (error bars represent ±1 SE) z-scores of audiovisual simultaneity threshold (top row) and point of subjective simultaneity (bottom row) for patients. Related to Figure 2 and Table 1. Panels (A) and (D) are for binocularly-deprived patients using the better eye; panels (B) and (E) are for monocularly-deprived patients using the deprived eye; panels (C) and (F) are for monocularly-deprived patients using the fellow eye.

Table S1. Clinical details of patients treated for binocular congenital cataracts. Related to Figure 1.

Patient	Age (years)	Diagnosis (days)	CL (days)[a]	Snellen acuity[b]	Stereo acuity (arc sec)[b,c]	Binocular fusion[b,d]	Eye tested	Nystagmus[e]	Additional details
B1	AV: 21.7 VT: 21.7	63	106	OD: 20/40 OS: 20/80 at 50 cm	>400	Suppresses OS	OD		Microcornea OU; Left membranectomy at 9 mo; Right pupilectomy & membranectomy at 2.5 yo; Glaucoma OU[1]
B2	AV: 31.8 VT: 33.0	142	187	OD: 20/25 OS: 20/63	>400	Diplopia	OD	Latent nystagmus OU	LET surgery at 1.7 yo; lens implant OU at 26 yo
B3	AV: 21.8 VT: 22.6	188	238	OD: 20/25 OS: 20/100	400	Fused	OD		Strab Sx OU at 9 mo; excision left orbital dermoid at 1.8 yo; Glaucoma OU[1]
B4	AV: 32.9 VT: 33.8	92	OD: 181 OS: 294	OD: 20/100 OS: 20/40	200	Alternator	OS	Nystagmus OU	RET Sx at 3 yo; Strab Sx OU at 5 & 10 yo
B5	AV: 14.2 VT: 14.2	0	17	OD: 20/32 OS: 20/32	140	Fused	OS		Glaucoma OU[1]
B6	AV: 20.3 VT: 21.2	61	92	OD: 20/63 OS: 20/32	>400	Diplopia	OS	Nystagmus OU	Microcornea OU; LET Sx at 1.5 yo; Strab Sx Ou at 3 yo; Glaucoma OU[1]
B7	VT: 34.2	0	129	OD: 20/50 OS: 20/80	>400	Alternator	OD		Capsular membrane needling OD at 3 mo; Secondary membranes & Elschnig's pearls removed OU at 9.7 mo
B8	AV: 21.3	12	48	OD: 20/32 OS: 20/40	>400	Fused	OD	Manifest Nystagmus OU	Glaucoma OU[1]
B9	AV: 22.1 VT: 22.1	104	134	OD: 20/125 OS: 20/100	>400	Diplopia	OS	Intermittent Nystagmus OU with an estotropia	Microcornea OU; Strab sx OU at 3.9 yo, Glaucoma OU[1]
B10	AV: 15.5 VT: 15.5	107	121	OD: 20/32 OS: 20/40	>400	Diplopia	OD	Nystagmus OU	Mild microcornea OU; Strab Sx OU at 1 yo
B11	AV: 22.2 VT: 22.9	113	147	OD: 20/32 OS: 20/32	>400	Diplopia	OD		Strab Sx OS at 1 yo

B12	AV: 19.1	1	34	OD: 20/40 OS: 20/63	>400	Diplopia	OD		Remove scar tissue at 2 mo; Strab Sx OU at 1 yo; Strab sx for LET at 3.8 yo; Glaucoma OS[1]
B13	AV: 18.9 VT: 19.7	0	9	OD: CF at 1.4m OS:20/32	>400	Suppresses OS	OS	Nystagmus OU	Ahmed Tube OD; Glaucoma OU[1]
B14	AV: 27.4 VT: 28.4	69	142	OD: 20/100 OS: 20/125	400	Diplopia	OD	Latent nystagmus	Glaucoma OU[1]; Glaucoma procedure at 4 yo; Hallermann Streiff Syndrome

AV: Audiovisual experiment; CF: counting fingers; CL: contact lens; LET: Left Esotropia; mo: months old; OD: Right eye; OS: Left eye; OU: Both eyes; RET: Right Esotropia; Sx: Surgery; Strab: Strabismus; VT: Visuotactile experiment; yo: years old.

[1] Patients with glaucoma had a cup/disc ratio < 0.7 in the tested eye, indicating little or no retinal damage from glaucoma

[a] Age at time of first optical correction after cataract surgery (defined as duration of deprivation).

[b] Measured at the time of test. Refractions are spherical equivalents.

[c] Stereo acuity was tested using Randot test.

[d] Binocular fusion was tested using Worth 4 Dot test.

[e] History of nystagmus since first optical correction.

Table S2. Clinical details of patients treated for monocular congenital cataract. Related to Figure 1.

Patient	Age (years)	Diagnosis (days)	CL (days)	Snellen acuity	Stereo acuity (arc sec)	Binocular fusion	Deprived eye	Nystagmus	Additional details
U1	AV: 17.0 VT: 17.7	0	41	OD: 20/200 OS: 20/20	>400	Suppresses OS	OD		tested right eye with +3.0 add
U2	AV: 28.1 VT: 29.1	150	245	OD: 20/16 OS: 20/200	>400	Diplopia	OS	Latent Nystagmus OU	Strab Sx OS at 1 yo, 1.7 yo and 25 yo
U3	AV: 11.4 VT: 11.9	0	26	OD: 20/20 OS: 20/32	>400	Suppresses OS	OS		
U4	AV: 16.0 VT: 16.2	0	18	OD: 20/20 OS: 20/63 at 50cm	>400	Suppreses OS	OS		Glaucoma OS[1]
U5	AV: 24.3 VT: 25.0	9	35	OD: 20/ 20 OS: 20/50	>400	Suppresses OS	OS	Latent Nystagmus	Strab Sx at 1.1 yo; membranectomy at 5.9 yo
U6	AV: 12.5 VT: 13.5	6	36	OD: 20/50 OS: 20/100	>400	Diplopia	OS		
U7	AV: 22.0 VT: 22.6	92	199	OD: 20/200 at 50 cm OS: 20/20	>400	Suppresses OD	OD		
U8	AV: 42.6	0	863	OD: 20/20 OS: HM at 50cm	>400	Suppresses OS	OS	Latent Nystagmus OS	Microcornea OS
U9	AV: 21.6	32	55	OD: 20/25 OS: 20/80	>400	Diplopia	OS		Strab Sx at 7 mo
U10	AV: 31.6	<90	124	OD: 20/50 OS: 20/25	>400	Diplopia	OD	Latent Nystagmus OU	Microcornea OD; Secondary membrane removal at 3.8 mo; RET Sx at 2.2 yo
U11	VT: 22.2	21	55	OD: 20/50 OS: 20/20	>400	Diplopia	OD	Latent Nystagmus OU	
U12	AV: 21.6 VT: 22.6	16	97	OD: 20/40 OS: 20/40	>400	Alternator	OS	Latent Nystagmus OS	Microcornea OS; tested left eye with +11 add since not wearing CL;

U13	AV: 30.3 VT: 31.3	75	173	OD: 20/12.5 OS: 20/160 at 2m	140	Suppresses OD	OS		Strab Sx OU at 4 yo; Strab Sx OS at 4 yo and 18 yo; Left goniotomy at 8 yo
U14	AV: 16.2 VT: 17.2	92	103	OD: 20/16 OS: 20/100	>400	Diplopia	OS		
U15	VT: 20.5	131	176	OD: 20/20 OS: 20/160 at 50cm	>400	Suppreses OS	OS		IOL at 4 yo

AV: Audiovisual experiment; CL: contact lens; HM: hand move; IOL: inter-ocular lens; mo: months old; OD: Right eye; OS: Left eye; OU: Both eyes; RET: Right Esotropia; Sx: Surgery; Strab: Strabismus; VT: Visuotactile experiment; yo: years old.

[1] Patients with glaucoma had a cup/disc ratio < 0.7 in the tested eye, indicating little or no retinal damage from glaucoma

Table S3. Correlations (R^2) between the point of subjective simultaneity (PSS) or the threshold of audiovisual simultaneity (δ) and (1) current visual acuity, (2) age, (3) duration of deprivation, and (4) the mean number of hours per day that the fellow eye had been patched for monocularly-deprived patients. Related to Table 1. None of the correlations was significant (all *p*s > .06). In summary, for both binocularly- and monocularly-deprived patients, neither the point of subjective simultaneity nor the threshold of audiovisual simultaneity correlated with visual acuity or age when tested. The latter suggests that the amount of lifetime perceptual experience (at least beyond 11 years of age – the youngest age tested) did not determine performance. Nor were these measures correlated with the duration of deprivation. Finally, for monocularly-deprived patients, the threshold of audiovisual simultaneity (δ) in the deprived eye and in the fellow eye did not correlate with the number of hours per day that the fellow eye had been patched.

	Binocular		Monocular			
	Better eye		Deprived eye		Fellow eye	
	PSS	δ	PSS	δ	PSS	δ
Current visual acuity	< 0.001	0.04	0.07	0.003	0.07	0.28
Age	0.21	0.22	0.17	0.19	0.19	0.06
Duration of Deprivation	0.26	0.11	0.15	0.13	0.24	0.02
Daily patch hours	n/a	n/a	0.18	0.03	0.28	0.002

Supplemental Experimental Procedures

Patients' history of visual deprivation and treatments

All patients included in the study were born with dense central cataracts in one or both eyes with no other abnormalities in the ocular media or the retina, and no evidence of persistent hyperplastic primary vitreous. However, patients who later developed common associated abnormalities such as strabismus, nystagmus, or microcornea were included, as well as those who developed glaucoma with no associalted retinal damage (i.e., a cup/disc ratio of less than 0.7).

Diagnosis of congenital cataracts was based on the first eye exam carried out within the first 6 months after birth. As in previous studies, we assumed that these patients had been visually deprived from birth because it would be rare to have dense cataracts develop rapidly between birth and 6 months of age [S1,S2]. Duration of deprivation was calculated as the period extending from birth until the age of first optical correction after surgery to remove the cataract. Even though there might be errors in diagnosis as we cannot be certain that all of these patients had dense central cataracts at birth, such errors in diagnosis are as likely to have occurred in binocularly- as in monocularly-deprived cases and thus would not explain any differences between the two groups of patients.

Patients treated for monocular congenital cataract were all instructed by their ophthalmologist to patch their fellow eye beginning shortly after the time of the first optical correction until at least 5 years of age. However, because of variation in instructions and compliance, the mean amount of patching from the time of the first optical correction until 5 years of age ranged from 2.5 to 7.1 waking hours per day (see [S3] for details of these calculations).

Data analysis

For both the audiovisual and visuotactile simultaneity judgments tasks, we calculated the percentage of simultaneous responses at each SOA for each participant. Individual data were fitted with the Matlab routine for simultaneity judgments task [S4]. This routine is based on an independent-channels model for stimulus timing judgments tasks proposed by García-Pérez and Alcalá-Quintana [S5]. In this model, it is assumed that the signals in each sensory modality (i.e., vision and audition, or vision and touch) are processed independently in each sensory pathway and then reach a central mechanism decoding their arrival time (i.e., perceived onset time of each signal). The difference of perceived onset time between the two signals, caused by the processing time and variance in each sensory pathway, would form a probability distribution rather than a constant. Participants' threshold of multisensory simultaneity is the criterion that determines whether or not they respond that the stimuli were simultaneous. Note that this model also assumes that the participants may make motor errors when responding.

The peripheral processing time of each sensory signal to reach a central mechanism is modeled by an exponential distribution determined by two parameters: the processing time (τ_i) and the processing variability (λ_i). In the audiovisual study, the processing time difference between two stimuli ($\tau = \tau A - \tau V$, so that a negative value indicates that the processing time was shorter for auditory than visual processing) and the processing variability of visual and auditory signal (λV and λA, respectively) were estimated. At a given SOA, a bilateral exponential distribution of the difference of perceived onset time of the two signals at the central mechanism was determined by λV, λA, and τ. A resolution parameter (δ), representing the threshold of simultaneity perception, was estimated such that a "simultaneous" response would be made when the difference of perceived onset time was smaller than the resolution parameter. Hence, the bigger the value of this parameter, the wider the audiovisual simultaneity window. In the model [S4,S5], the resolution parameter (δ) corresponds to half of the width of the simultaneity window when the criterion was set at 50% simultaneous response either on the auditory leading or the visual leading side. The point of subjective simultaneity (PSS) is the midpoint of the audiovisual simultaneity window determined by the resolution parameter (δ). Finally, the response errors when a participant misreported "simultaneous" in the auditory-leading trials (εAF) and in the visual-leading trials (εVF), as well as the "not simultaneous" errors in the 0-ms trials (εS) were estimated. Note that these response errors include both participants' *lapse* due to blink or inattention as well as their mistakes of motor responding, since these two types of errors are hard to distinguish [S6]. The starting values used to fit the data in the present study were as follows [S4]: LamBounds = [1/500 1/1]; TauBounds = [-Inf Inf]; DeltaBounds = [0 Inf]; LamTStart = [1/70 1/10]; LamRStart = [1/70 1/10]; TauStart = [-70 70]; DeltaStart = [20 150]; ErrStart = [0.05]; Model = 1; SampleSize = 1500; ConfCoef = 95; FixedSeed = true.

In the visuotactile study, the data analysis was exactly the same, with all of the parameters associated with auditory processing now associated with tactile processing.

Supplemental References

S1. Ellemberg, D., Lewis, T. L., Maurer, D., and Brent, H. P. (2000). Influence of monocular deprivation during infancy on the later development of spatial and temporal vision. Vision Res. *40*, 3283–3295.

S2. Ellemberg, D., Lewis, T. L., Maurer, D., Brar, S., and Brent, H. P. (2002). Better perception of global motion after monocular than after binocular deprivation. Vision Res. *42*, 169–179.

S3. Lewis, T. L., Maurer, D., and Brent, H. P. (1995). Development of grating acuity in children treated for unilateral or bilateral congenital cataract. Invest. Ophth. Vis. Sci. *36,* 2080–2095.
S4. Alcalá-Quintana, R., and García-Pérez, M. A. (2013). Fitting model-based psychometric functions to simultaneity and temporal-order judgment data: MATLAB and R routines. Behav. Res. Methods *45*, 972–998.
S5. García-Pérez, M. A., and Alcalá-Quintana, R. (2012a). On the discrepant results in synchrony judgment and temporal-order judgment tasks: A quantitative model. Psychon. B. Rev. *19*, 820–846.
S6. García-Pérez, M. A., and Alcalá-Quintana, R. (2012b). Response errors explain the failure of independent-channels models of perception of temporal order. Front. Psychol. *3:94*.