Report

Current Biology

Development of multisensory integration following prolonged early-onset visual deprivation

Highlights

- Congenitally blind individuals regaining sight late acquire multisensory integration abilities
- The ability to integrate vision and touch develops quickly after surgery
- Some individuals reach optimal integration levels within years to benefit perception
- The development is based on experience and depends on post-surgical visual acuity

Authors

Irene Senna, Elena Andres, Ayelet McKyton, Itay Ben-Zion, Ehud Zohary, Marc O. Ernst

Correspondence

irene.senna@uni-ulm.de (I.S.),
udiz@mail.huji.ac.il (E.Z.)

In brief

In healthy individuals, multisensory integration takes years to mature. Senna et al. demonstrate that early exposure to multisensory inputs is not essential for such development. Congenitally blind people develop integration of vision and touch after sight-restoration surgery late in life. Developing integration is crucial for adept behavior.

Senna et al., 2021, Current Biology *31*, 4879–4885
November 8, 2021 © 2021 Elsevier Inc.
https://doi.org/10.1016/j.cub.2021.08.060

Current Biology

CellPress

Report

Development of multisensory integration following prolonged early-onset visual deprivation

Irene Senna,[1,6,7,*] Elena Andres,[2,6] Ayelet McKyton,[3] Itay Ben-Zion,[4] Ehud Zohary,[2,5,*] and Marc O. Ernst[1]
[1]Applied Cognitive Psychology, Faculty for Computer Science, Engineering, and Psychology, Ulm University, Ulm 89081, Germany
[2]The Edmond & Lily Safra Center for Brain Science, The Hebrew University of Jerusalem, Jerusalem 91904, Israel
[3]fMRI Unit, Department of Neurology, Hadassah Medical Organization and Faculty of Medicine, Hebrew University of Jerusalem, Jerusalem 91120, Israel
[4]Pediatric Ophthalmology Service, Padeh Medical Center, Tiberias, Israel
[5]Neurobiology Department, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel
[6]These authors contributed equally
[7]Lead contact
*Correspondence: irene.senna@uni-ulm.de (I.S.), udiz@mail.huji.ac.il (E.Z.)
https://doi.org/10.1016/j.cub.2021.08.060

SUMMARY

Adult humans make effortless use of multisensory signals and typically integrate them in an optimal fashion.[1] This remarkable ability takes many years for normally sighted children to develop.[2,3] Would individuals born blind or with extremely low vision still be able to develop multisensory integration later in life when surgically treated for sight restoration? Late acquisition of such capability would be a vivid example of the brain's ability to retain high levels of plasticity. We studied the development of multisensory integration in individuals suffering from congenital dense bilateral cataract, surgically treated years after birth. We assessed cataract-treated individuals' reliance on their restored visual abilities when estimating the size of an object simultaneously explored by touch. Within weeks to months after surgery, when combining information from vision and touch, they developed a multisensory weighting behavior similar to matched typically sighted controls. Next, we tested whether cataract-treated individuals benefited from integrating vision with touch by increasing the precision of size estimates, as it occurs when integrating signals in a statistically optimal fashion.[1] For participants retested multiple times, such a benefit developed within months after surgery to levels of precision indistinguishable from optimal behavior. To summarize, the development of multisensory integration does not merely depend on age, but requires extensive multisensory experience with the world, rendered possible by the improved post-surgical visual acuity. We conclude that early exposure to multisensory signals is not essential for the development of multisensory integration, which can still be acquired even after many years of visual deprivation.

RESULTS AND DISCUSSION

We investigated whether young individuals suffering from congenital dense bilateral cataracts for many years are able to efficiently combine information from different senses, once their vision is surgically restored. Adults typically integrate multisensory inputs derived from a given object property (e.g., an object's size estimated by vision and touch) in a statistically optimal fashion. That is, they weigh information from each modality according to its reliability (inverse variance), thereby achieving the most precise estimate possible.[1] While some interaction between the senses occurs early in life,[4–14] optimal multisensory integration is achieved only at 8–10 years.[2,3,15–18] This long developmental path may be due to the inherent computational complexity: one must first infer which signals arise from a common source,[18] and know the mapping between the signals, before integrating them into a multisensory estimate. Learning such mapping, for example, knowing which retinal image-size corresponds to which postural aperture between fingers holding an object, likely requires extensive experience.[19–25] Thus, one may wonder whether early experience with multisensory input is necessary for optimal integration to develop or whether integration can also be acquired later in life. Therefore, we surgically treated and tested Ethiopian participants suffering from bilateral cataract, which rendered them blind (or with severe low vision) for years (Figure S1A; Table S1). In experiment 1, we investigated participants' reliance on both modalities when estimating object size visual-haptically. In experiment 2, we assessed whether jointly using both modalities yields the predicted benefit of increased multisensory precision.[1]

Cataract-treated participants learn to weigh information from vision and touch

We tested 16 cataract-treated individuals who had eye surgery a few days to several years earlier. They were asked to match the size of an object (standard) presented visually (V), haptically (H), or visual-haptically (VH) to 1 of 10 objects varying in size (comparison), explored either visually or haptically.[14,26–28] Thus, there

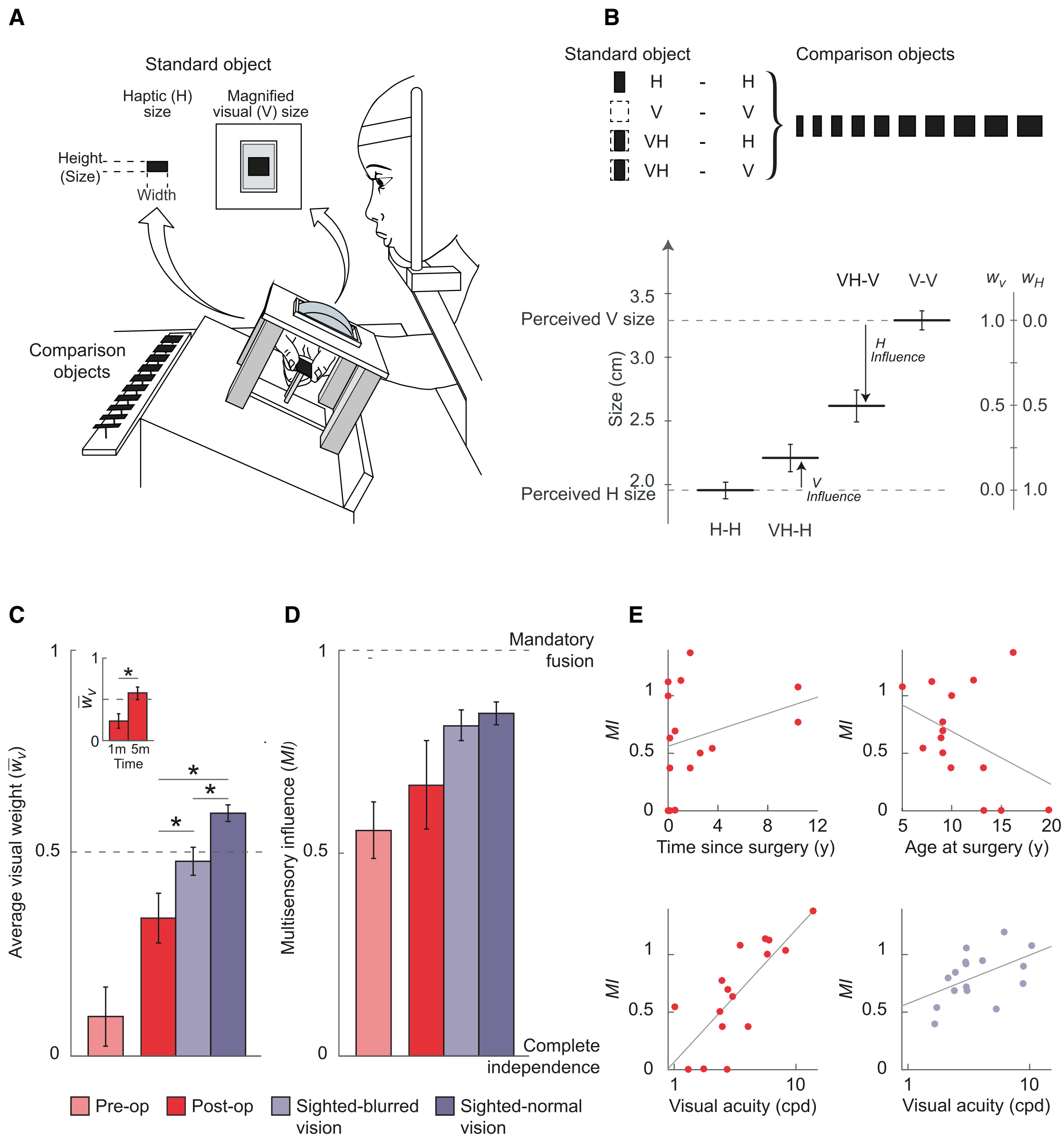

Figure 1. Multisensory weighting behavior

(A) Apparatus.

(B) Size estimation: participants haptically (H), visually (V), and visual-haptically (VH) explored a standard object (2 cm high), whose visual shape was magnified along its height via a distortion lens (magnification = 1.75). They matched the perceived size of the standard to a set of 10 comparison objects differing in size (hidden from view during standard exploration). The comparison was presented either visually (V) or haptically (H), yielding 2 unisensory and 2 multisensory conditions: V-V, H-H, VH-H, and VH-V (upper panel). The lower panel shows the mean of the selected comparisons' sizes for the cataract-treated individuals (n = 16). The lower and upper dashed lines indicate perceived haptic and visual size in H-H and V-V conditions, respectively. Results from all 4 conditions differed from one another (Wilcoxon signed-rank tests, all $p \leq 0.015$, following Friedman test, $\chi^2_{(3,45)} = 34.5$, $p < 0.0001$, all p values of pairwise comparisons in the article are Bonferroni-Holm corrected). The fact that the visual-haptic size estimates in the multisensory conditions (VH-H, VH-V) lie between visual (V-V) and haptic (H-H) estimates indicates multisensory integration. Integration took place on a trial-by-trial basis and was not the average of a switching strategy between the senses (Figure S2). The arrows indicate the direction of visual and haptic influence on the multisensory estimates. The scale for visual and haptic weights (w_V, w_H) for the multisensory conditions is shown to the right.

(C) Average visual weight $\overline{w}_V$ across VH-V and VH-H in pre-operated cataract participants (n = 4), post-operated cataract participants (n = 16), and sighted controls either tested in normal visual conditions (n = 112) or with visual blur (n = 16). Cataract-treated participants weighted vision less than sighted controls tested with or without visual blur (Wilcoxon rank-sum test, p = 0.04 and p = 0.0008, respectively, following Kruskal-Wallis test, $\chi^2_{(2)} = 17.3$, p = 0.0002). Importantly, cataract-treated participants' $\overline{w}_V$ was >0 (1-sample Wilcoxon signed-rank test, p = 0.0006). The inset shows the development of $\overline{w}_V$ in a subset of 7 cataract-treated participants tested twice: at <1 month and at 5 months after surgery, on average. In the follow-up, their $\overline{w}_V$ was indistinguishable from controls.

(D) Multisensory influence (*MI*) between vision and haptics. Mandatory fusion: *MI* = 1; independence: *MI* = 0. *MI* did not differ across groups (Kruskal-Wallis test, $\chi^2_{(2)} = 2.15$, p = 0.33).

(legend continued on next page)

were 2 unisensory and 2 multisensory conditions: in H-H, standard and comparison were both presented haptically; in V-V, standard and comparison were both presented visually; in VH-V, the standard was presented visual-haptically, the comparison visually; and in VH-H, the standard was presented visual-haptically, the comparison haptically. To assess the influence of vision on touch, we introduced a discrepancy between the two senses by vertically magnifying the seen height of the standard using a distortion lens (Figure 1A). The unisensory conditions (V-V, H-H) served as a baseline for the size estimates derived from vision or touch alone. They revealed that cataract-treated individuals were able to use their restored vision, besides touch, for performing the task (Figure 1B). The multisensory conditions (VH-V, VH-H) were critical for assessing participants' reliance on each modality when making multisensory size estimates. If the cataract-treated individuals fail to use visual information, then their multisensory estimates should not differ from their unisensory haptic estimate (H-H). In contrast, if they are able to integrate visual and haptic information (i.e., they rely also on vision, thereby giving weight to both signals), then the visual-haptic estimate should be between that of vision (V-V) and touch (H-H).[26–28] This is what we found in both multisensory conditions (VH-V, VH-H): the multisensory estimates were between the unisensory visual and haptic estimates (Figure 1B). To quantify the reliance on vision for multisensory size estimation, we determined the average weight given to vision $\overline{w}_V$ across VH-H and VH-V (Equation 2, STAR Methods). This weight is evaluated against that of 2 control groups (Figure 1C): the first group consisted of 112 typically developing sighted individuals in a similar age range as the cataract treated. To exclude that differences in performance between the groups simply resulted from the poorer visual acuity of the cataract treated even after surgery,[29,30] we blurred the visual stimulus in a second age-matched control group, individually matching controls' visual acuity to that of the cataract-treated participants (n = 16). Acuity was assessed by measuring the contrast sensitivity function[31] (STAR Methods). Before surgery, most of the participants were unable to perform the visual task, and the 4 who showed some performance had a $\overline{w}_V$ close to 0. After surgery, participants' average visual weight was $\overline{w}_V$ = .34, which was significantly greater than 0, showing that they made use of their improved vision for multisensory estimates. Still, overall, the cataract-treated group weighted vision less than both control groups (Figure 1C). Importantly, we were able to assess a subset of cataract-treated individuals twice (n = 7): shortly after surgery and again 4 months later. Even though their $\overline{w}_V$ differed from 0 already in the first test ($p = 0.047$), it increased significantly in the retest (Wilcoxon signed-rank test, $p = 0.016$; Figure 1C, inset), and was indistinguishable from controls ($p = 0.87$). While the average visual weight across both multisensory conditions provides a general estimate of the use of vision when judging multisensory stimuli, with the current task that ensured a brief experiment, we are unable to assess whether $\overline{w}_V$ is set to the theoretical optimum according to Equation 7. This is because using this task does not allow assessing the precision σ_i of the individual visual and haptic estimates (although they are investigated in experiment 2).

Importantly, however, the advantage of the present task is that it enables us to assess how much individuals actually integrate vision with touch and to follow the developmental path of multisensory integration. This is achieved by investigating the possible difference between the two multisensory size estimates obtained from VH-V and VH-H. Depending on which modality is used for comparison, these conditions measure two different multisensory influences; comparing the multisensory standard to a visual comparison (VH-V) measures the influence of touch on vision, while comparing the multisensory standard to a haptic comparison (VH-H) measures the influence of vision on touch ("haptic influence" versus "visual influence," respectively, in Figure 1B).[27] For illustration, consider the case that vision exerts no influence on touch; then, adding vision to the multisensory stimulus in the VH-H condition would not lead to a size estimation different from that in H-H (i.e., participants rely on touch only, therefore the visual weight $w_V = 0$). Similarly, when touch exerts no influence on vision, adding touch to the multisensory stimulus in the VH-V condition would not lead to a size estimation different from that in V-V (i.e., participants rely on vision only, thus $w_V = 1$). This would happen if the multisensory signals were not integrated but treated as independent.[27] However, even if the signals were independent, the resulting average visual weight across the two multisensory conditions would be $\overline{w}_V = 0.5$ ($w_V = 0$ in one condition and 1 in the other). As many other combinations of visual weights across the two multisensory conditions can also result in an average visual weight of $\overline{w}_V = 0.5$ (e.g., 0.4/0.6, 0.3/0.7), this average weight is clearly not able to capture the degree of multisensory integration as measured by the mutual influence (*MI*) of vision on touch and vice versa. Unlike independence, in which the multisensory signals are kept separate with no mutual influence, in the case of mandatory fusion vision and touch influence each other to a degree that they are merged into a unique estimate of visual and haptic size. That is, in mandatory fusion, both VH-V and VH-H size estimates agree, and hence their visual weights are identical, with no difference.[27] Learning to integrate multisensory signals (as, for example, during development) should follow a trajectory along a continuum from initial independence to near mandatory fusion.[19,27,32–34] Can we trace such a trajectory in the cataract-treated individuals?

Overall, cataract-treated individuals show an intermediate result between independence and mandatory fusion (Figure 1B). Vision was able to influence touch ("visual influence" in VH-H) and touch was able to influence vision ("haptic influence" in VH-V), but the two effects did not add up to an identical size estimate. We quantified such *MI* (visual influence on touch in VH-H and haptic influence on vision in VH-V) as the sum of visual and haptic weights from those two conditions: $MI = w_V(\text{VH-H}) + w_H(\text{VH-V})$ (Equations 3, 4, and 5). Defined in this way, $MI = 0$ means no influence (i.e., independence), while $MI = 1$ corresponds to mandatory fusion. *MI* can thus be used to monitor the development of multisensory integration. On average, for

(E) Correlations between *MI* and time since surgery, age at surgery, and visual acuity in cataract-treated (red circles, cf. Figure S1): *MI* was higher at higher visual acuity, longer time since surgery, and younger age at surgery. *MI* also tended to correlate with visual acuity in blurred-vision controls ($r = 0.42$, $p = 0.1$, lilac circles). Error bars represent SEMs across participants.
See also Figure S1 and S2 and Table S1.

the cataract-treated individuals, $MI = 0.67 \pm 0.11$ (Figure 1D), which is well above 0 (Wilcoxon signed-rank test, $p = 0.0001$). When comparing the cataract treated to both control groups, we observed an overall smaller *MI*, although the 3 groups did not differ significantly. Interestingly, also in controls, *MI* was <1 (with blur: 0.81 ± 0.05; without blur: 0.84 ± 0.03), indicating that also controls did not mandatorily fuse the signals. This is in agreement with previous studies; while healthy adults make optimal use of multisensory signals—i.e., maximizing the precision of multisensory estimates—they typically do not mandatorily fuse those signals.[26,27,35,36] This seems to happen to retain some access to the input signals during multisensory integration (e.g., being able to detect a discrepancy between the multisensory signals seems necessary for cross-modal calibration[18,19,36]).

Although not statistically significant, it seems that the four individuals who could perform the visual task before surgery show also some integration abilities (with a low visual weight; Figures 1C and 1D).

The average *MI*s include all cataract-treated individuals with different ages, tested at different times after surgery, and varying in post-surgical visual acuity. To unravel the large variability in this heterogeneous group and to investigate which factors affected performance at the individual level, we ran a multiple regression analysis on *MI* considering participants' age at test, time since surgery, and visual acuity (Figure 1E). *MI* was higher with better visual acuity ($t = 5.83$, $p = 0.0003$) and longer time since surgery ($t = 3.44$, $p = 0.011$). This finding indicates that the multisensory influence (thus, multisensory integration) increased with post-surgical experience, especially in participants benefiting from better-quality visual and multisensory input due to higher visual acuity. Time since surgery and visual acuity did not significantly correlate ($r = -0.12$, $p = 0.67$), meaning that the development of integration with experience after surgery is not just a spurious correlation. There was a negative trend showing decreasing *MI* with age at test ($t = 2.17$, $p = 0.053$). As age at test and age at surgery were correlated, the result is independent of which measure is used for the multiple regression (Figure 1E). Sighted participants showed the opposite trend: higher *MI* with increasing age (Figure S1B). This finding suggests that participants operated on later in life, thus lacking good-quality visual and multisensory experience for longer, are less likely to integrate multisensory inputs. These findings show that cataract-treated individuals develop the ability to weight visual and haptic inputs for integrating size estimates, approaching the level of matched sighted controls within a few months after surgery.

Cataract-treated participants learn to benefit from multisensory integration

Sighted adults typically integrate redundant multisensory information in a statistically optimal fashion according to "maximum-likelihood estimation" (MLE).[1] By integrating information optimally, the multisensory estimate is more precise than each unisensory estimate alone (Equations 6, 7, and 8). Hence, a critical test for efficient multisensory integration requires showing increasing precision, in comparison to both unisensory estimates. Thus, in experiment 2, we compared the precision of visual, haptic, and visual-haptic size judgments.

A total of 23 cataract-treated participants discriminated object size using vision, touch, or both. In each trial, they consecutively explored a standard cube (50 cm) and a comparison cube (varying in size across trials) and reported which was bigger. The proportion of "comparison bigger" responses against size was fitted with a cumulative Gaussian[37] per condition (visual, haptic, visual-haptic) and participant. From these fits, we determined estimate precision as the size-discrimination threshold, the just noticeable differences (JNDs). If participants benefit from multisensory integration, the visual-haptic JND_{VH} should be lower than either the visual JND_V or the haptic JND_H. However, the group's JND_{VH} was not smaller than the better of the two unisensory JNDs (typically JND_H; Figure 2A) and it was worse than the maximum likelihood estimation (MLE)-predicted optimal visual-haptic $JND_{optimal}$ (Equation 9, Wilcoxon signed-rank test, $p = 0.015$). Hence, overall, this heterogenous group did not optimally integrate visual-haptic information.

However, this group also included participants who were operated on only days pre-testing. To unravel the development of multisensory integration, we analyzed whether multisensory precision improved with time after surgery, independent of age or visual acuity. For each participant, we quantified the improvement in precision as *integration gain* $g_{integration}$, the log-transformed ratio between the better unisensory JND (usually haptics) and the visual-haptic JND_{VH} (Equation 10; Figure 2B). Typically, learning follows an exponential function—here, with a fixed asymptote c at the MLE-predicted *optimal gain* $\overline{g}_{optimal}$ for the group mean (STAR Methods). $g_{integration} < 0$ means a decrement in multisensory performance relative to the better unisensory estimate (meaning that vision is used but interferes). $g_{integration} > 0$ indicates a benefit from integration, with the theoretical maximum at the optimal gain. To assess the developmental path, we fitted an exponential to the multisensory gain data as a function of time since surgery, with parameters amplitude a and time constant $-b$:

$$g_{integration} = ae^{-bx} + c \quad \text{(Equation 1)}$$

$g_{integration}$ improved over the first months after surgery (Figure 2B). However, the fitted learning rate $b = 0.35$ was not significantly different from 0 (95% confidence interval [CI] for b: $[-0.4, 1.1]$). Thus, the exact time at which integration reached optimal performance levels could not be determined with certainty from the group data (Figure 2B, inset, difference between observed and predicted optimal integration gains): while some participants showed optimal performance a few months after surgery, others did not. This is likely due to the large intersubject variability caused by other factors. One of those factors is age, which correlated with $g_{integration}$ (Pearson $r = 0.47$, $p = 0.024$), as color-coded in Figure 2B, with older participants closer to optimal integration performance. This finding indicates that both age and post-surgical experience contribute to the development of optimal multisensory integration. $g_{integration}$ relates to visual acuity (thus JND_V) in a non-linear v-shaped function (Equation 10). Therefore, as expected, visual acuity did not significantly correlate with $g_{integration}$(Figure S1C).

We had the opportunity to test 5 cataract-treated participants multiple times. This allowed assessing the effects of post-surgical visual and multisensory experience, while controlling for the other sources of intersubject variability. A few days after surgery, participants were far from showing optimal behavior (i.e., their JND_{VH} was higher than the MLE-predicted $JND_{optimal}$, 1-tailed,

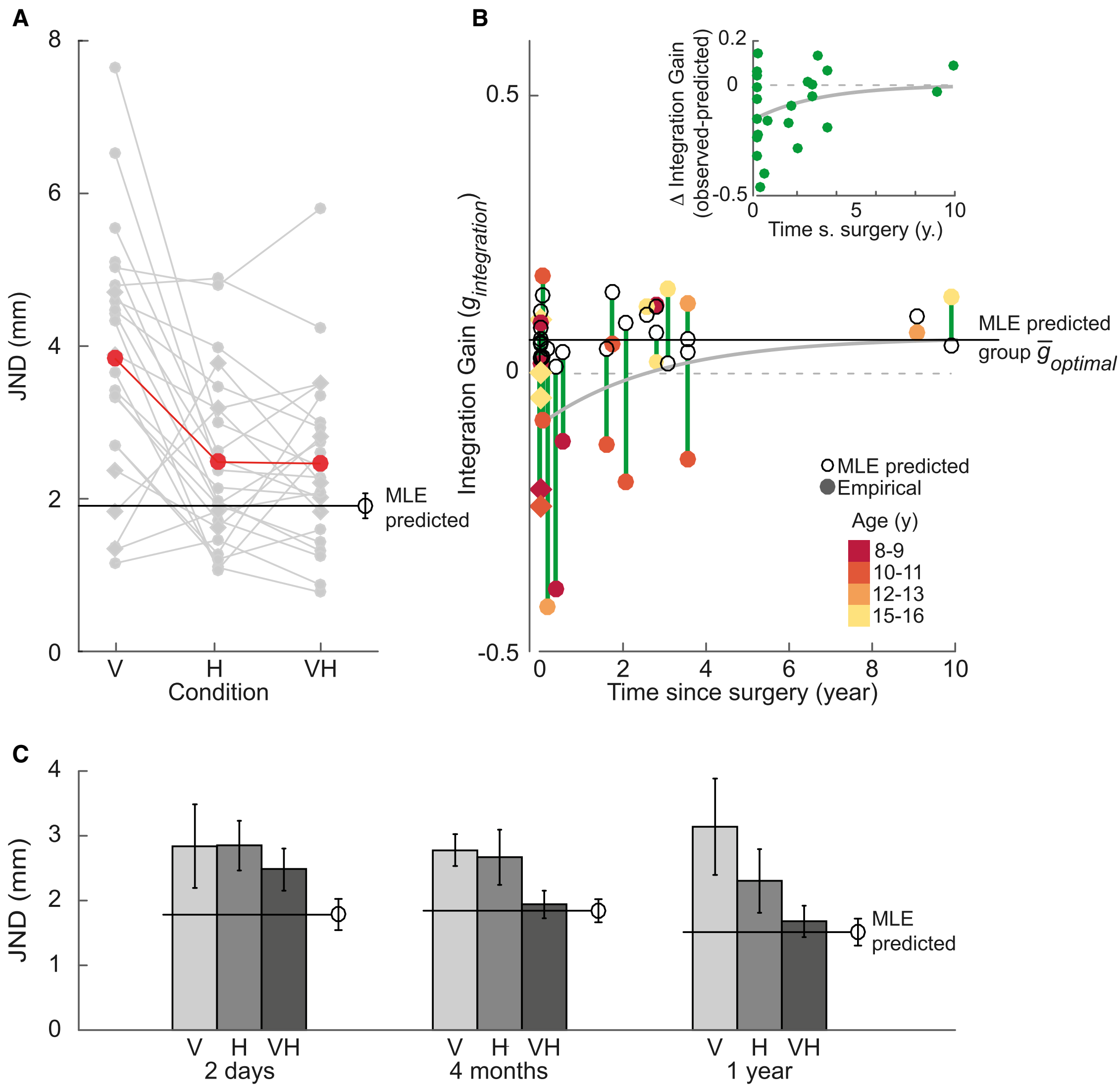

Figure 2. Multisensory discrimination performance

(A) JNDs for V, H, and VH size discrimination in cataract-treated individuals in the first post-surgical test (n = 23). Red circles indicate group average performance for each condition, while individual JNDs are shown in gray circles connected by lines. The average multisensory performance (JND_{VH}) was not better than the unisensory haptic JND_H (p = 0.72, Wilcoxon signed-rank test following Friedman test $\chi^2_{(2)} = 10.75$, p = 0.0046), and was worse than MLE-predicted optimal performance (open circle: $JND_{optimal}$ averaged across participants). Although participants' visual acuity improved after surgery (Figure S1A), leading to a better visual size discrimination performance, especially in participants with higher post-surgical visual acuity (Figure S1D), their visual JND_V was still worse than their JND_H or JND_{VH} (both p = 0.0045).

(B) Integration gain $g_{integration}$ as a function of time since surgery in the first post-surgical test (filled circles, with brighter colors indicating older individuals). An advantage of the multisensory over the unisensory conditions leads to positive values. The dashed line indicates no difference between multisensory and the better unisensory performance. The black open circles represent the MLE-predicted optimal gain $g_{optimal}$ for each participant. The development is indicated by an exponential fit to $g_{integration}$ (solid gray line). The curve's asymptote c is fixed to the group average optimal gain $\bar{g}_{optimal}$ (solid black line). The individual difference between predicted and observed integration gains ($\Delta g_{integration}$) is drawn as green connecting lines. For better visibility of the development of this difference over time, the inset represents the exponential time course across participants (green dots), with the dashed line indicating optimality. Diamonds in (A) and (B) indicate participants who were tested in follow-up sessions (see C).

(C) Mean V, H, and VH JNDs for a subset of 5 participants tested multiple times after surgery (the first post-surgical test, after 2 days, contributes also to Figures 1A and 1B, diamonds). The black open circle to the right of each graph indicates the group mean of the MLE-predicted $JND_{optimal}$. Error bars represent SEMs across participants.

See also Figure S1 and Table S1.

p = 0.04). Four months to 1 year later, the precision of their visual-haptic estimate did not differ from the MLE prediction anymore (p = 0.4 and p = 0.35, respectively; Figure 2C), indicating that this subset of participants quickly learned to make efficient use of multisensory information, showing behavior indistinguishable from the optimal.

To summarize, both experiments provide converging evidence that, despite prolonged early-onset visual deprivation, late cataract-treated individuals develop multisensory integration abilities within months from surgery. Although visual acuity substantially improves immediately after cataract removal (usually ameliorating for 6 months post-surgery),[38–40] it still lags behind the normative range. However, despite their often still-poor post-surgical visual acuity, cataract-treated participants frequently integrate visual and haptic information at a level comparable to sighted controls (matched for acuity), in some cases even showing performance indistinguishable from optimality. Previous studies demonstrated that a few months of early visual deprivation are sufficient to permanently alter multisensory interactions at both neural and behavioral levels.[41–46] However, our data suggest that prolonged early-onset visual deprivation does not destroy the ability to acquire multisensory integration following surgery (cf. Putzar et al.[47]). This finding is in accordance with recent observations showing that late-treated individuals can correctly match the shape of a haptically[48] or visual-haptically[49] explored object to its visual counterpart within days after surgery. It is likely that simple forms of cross-modal interaction (e.g., cross-modal matching) develop within days,[48,49] while more complex forms of integration may require months to years.

While healthy individuals reach adult-like size integration performance only at ages 8–10,[2,3,15–18,28] our findings show that cataract-treated participants in several cases (not all) require much shorter periods of visual and multisensory experience to develop optimal multisensory integration. The ability to dynamically adapt and learn new mappings within and between sensory modalities is a key feature of the human brain, which is able to quickly pick up new statistical regularities from the environment.[20–24] Here, we show that this crucial feature is not lost even after years of visual impairment, suggesting that the period for the development of multisensory integration extends well beyond early childhood.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Participants
- METHOD DETAILS
  - Experiment 1
  - Experiment 2
  - Assessment of the spatial visual acuity
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Experiment 1
  - Experiment 2

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2021.08.060.

ACKNOWLEDGMENTS

We thank Zemene Zeleke for coordinating and organizing our testing in Ethiopia. We thank Nicola Schlesinger and Angelika Schymonski for their help in data collection in Germany. This research was supported by the Deutsche Forschungsgemeinschaft (DFG) DIP-Grant awarded to E.Z. (ZO 349/1-1) and M.O.E. (ER 542/3-1).

AUTHOR CONTRIBUTIONS

Conceptualization, M.O.E. and I.S.; methodology, I.S., E.A., E.Z., and M.O.E.; software, I.S. and A.M.; investigation, I.S., E.A., A.M., M.O.E., and I.B.-Z.; formal analyses, I.S., E.A., and M.O.E.; resources, E.Z. and I.B.-Z.; data curation, I.S., E.A., and A.M.; writing – original draft, I.S.; writing – review & editing, I.S., M.O.E., E.Z., E.A., and A.M.; supervision, M.O.E. and E.Z.; funding acquisition, M.O.E. and E.Z.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 16, 2021
Revised: July 12, 2021
Accepted: August 23, 2021
Published: September 16, 2021

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. Gori, M., Del Viva, M., Sandini, G., and Burr, D.C. (2008). Young children do not integrate visual and haptic form information. Curr. Biol. *18*, 694–698.
3. Nardini, M., Jones, P., Bedford, R., and Braddick, O. (2008). Development of cue integration in human navigation. Curr. Biol. *18*, 689–693.
4. Gottfried, A.W., Rose, S.A., and Bridger, W.H. (1977). Cross-modal transfer in human infants. Child Dev. *48*, 118–123.
5. Spelke, E.S. (1979). Perceiving bimodally specified events in infancy. Dev. Psychol. *15*, 626–636.
6. Rose, S.A., and Ruff, H.A. (1987). Cross-modal abilities in human infants. In Handbook of Infant Development, J.D. Osofsky, ed. (Wiley), pp. 318–362.
7. Lewkowicz, D.J. (1992). Infants' responsiveness to the auditory and visual attributes of a sounding/moving stimulus. Percept. Psychophys. *52*, 519–528.
8. Lewkowicz, D.J. (1996). Perception of auditory-visual temporal synchrony in human infants. J. Exp. Psychol. Hum. Percept. Perform. *22*, 1094–1106.
9. Rosenblum, L.D., Schmuckler, M.A., and Johnson, J.A. (1997). The McGurk effect in infants. Percept. Psychophys. *59*, 347–357.
10. Bahrick, L.E., and Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. Dev. Psychol. *36*, 190–201.
11. Bahrick, L.E. (2001). Increasing specificity in perceptual development: infants' detection of nested levels of multimodal stimulation. J. Exp. Child Psychol. *79*, 253–270.
12. Scheier, C., Lewkowicz, D.J., and Shimojo, S. (2003). Sound induces perceptual reorganization of an ambiguous motion display in human infants. Dev. Sci. *6*, 233–241.
13. Walker, P., Bremner, J.G., Mason, U., Spring, J., Mattock, K., Slater, A., and Johnson, S.P. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. Psychol. Sci. *21*, 21–25.
14. McGurk, H.P., and Roderick, P. (1980). Intermodal coordination in young children: vision and touch. Dev. Psychol. *16*, 679–680.
15. Gori, M., Sandini, G., and Burr, D. (2012). Development of visuo-auditory integration in space and time. Front. Integr. Nuerosci. *6*, 77.

16. Petrini, K., Remark, A., Smith, L., and Nardini, M. (2014). When vision is not an option: children's integration of auditory and haptic information is suboptimal. Dev. Sci. *17*, 376–387.
17. 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.
18. Ernst, M.O. (2008). Multisensory integration: a late bloomer. Curr. Biol. *18*, R519–R521.
19. Ernst, M.O. (2007). Learning to integrate arbitrary signals from vision and touch. J. Vis. *7*, 7.1–7.14.
20. Epstein, W., and Morgan, C.L. (1970). Adaptation to uniocular image magnification: modification of the disparity-depth relationship. Am. J. Psychol. *83*, 322–329.
21. Ernst, M.O., Banks, M.S., and Bülthoff, H.H. (2000). Touch can change visual slant perception. Nat. Neurosci. *3*, 69–73.
22. Adams, W.J., Banks, M.S., and van Ee, R. (2001). Adaptation to three-dimensional distortions in human vision. Nat. Neurosci. *4*, 1063–1064.
23. Atkins, J.E., Fiser, J., and Jacobs, R.A. (2001). Experience-dependent visual cue integration based on consistencies between visual and haptic percepts. Vision Res. *41*, 449–461.
24. Parise, C.V., Knorre, K., and Ernst, M.O. (2014). Natural auditory scene statistics shapes human spatial hearing. Proc. Natl. Acad. Sci. USA *111*, 6104–6108.
25. Scheller, M., Proulx, M.J., de Haan, M., Dahlmann-Noor, A., and Petrini, K. (2021). Late- but not early-onset blindness impairs the development of audio-haptic multisensory integration. Dev. Sci. *24*, e13001.
26. Rock, I., and Victor, J. (1964). Vision and touch: an experimentally created conflict between the two senses. Science *143*, 594–596.
27. Helbig, H.B., and Ernst, M.O. (2007). Knowledge about a common source can promote visual-haptic integration. Perception *36*, 1523–1533.
28. Jovanovic, B., and Drewing, K. (2014). The influence of intersensory discrepancy on visuo-haptic integration is similar in 6-year-old children and adults. Front. Psychol. *5*, 57.
29. Lewis, T.L., and Maurer, D. (2005). Multiple sensitive periods in human visual development: evidence from visually deprived children. Dev. Psychobiol. *46*, 163–183.
30. Maurer, D., Ellemberg, D., and Lewis, T.L. (2006). Repeated measurements of contrast sensitivity reveal limits to visual plasticity after early binocular deprivation in humans. Neuropsychologia *44*, 2104–2112.
31. McKyton, A., Ben-Zion, I., and Zohary, E. (2018). Lack of automatic imitation in newly sighted individuals. Psychol. Sci. *29*, 304–310.
32. Ernst, M.O. (2005). A Bayesian view on multimodal cue integration. In Human Body Perception from the Inside Out, G. Knoblich, I.M. Thornton, M. Grosjean, and M. Shiffrar, eds. (Oxford University Press).
33. Bresciani, J.-P., Dammeier, F., and Ernst, M.O. (2006). Vision and touch are automatically integrated for the perception of sequences of events. J. Vis. *6*, 554–564.
34. Ernst, M.O. (2012). Optimal multisensory integration: assumptions and limits. In The New Handbook of Multisensory Processes, B.E. Stein, ed. (MIT Press), pp. 1084–1124.
35. Hershberger, W.A., and Misceo, G.F. (1996). Touch dominates haptic estimates of discordant visual-haptic size. Percept. Psychophys. *58*, 1124–1132.
36. Hillis, J.M., Ernst, M.O., Banks, M.S., and Landy, M.S. (2002). Combining sensory information: mandatory fusion within, but not between, senses. Science *298*, 1627–1630.
37. Wichmann, F.A., and Hill, N.J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. Percept. Psychophys. *63*, 1293–1313.
38. Ganesh, S., Arora, P., Sethi, S., Gandhi, T.K., Kalia, A., Chatterjee, G., and Sinha, P. (2014). Results of late surgical intervention in children with early-onset bilateral cataracts. Br. J. Ophthalmol. *98*, 1424–1428.
39. Kalia, A., Lesmes, L.A., Dorr, M., Gandhi, T., Chatterjee, G., Ganesh, S., Bex, P.J., and Sinha, P. (2014). Development of pattern vision following early and extended blindness. Proc. Natl. Acad. Sci. USA *111*, 2035–2039.
40. Orlov, T., Raveh, M., McKyton, A., Ben-Zion, I., and Zohary, E. (2021). Learning to perceive shape from temporal integration following late emergence from blindness. Curr. Biol. *31*, 3162–3167.e5.
41. 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.
42. Putzar, L., Goerendt, I., Heed, T., Richard, G., Büchel, C., and Röder, B. (2010). The neural basis of lip-reading capabilities is altered by early visual deprivation. Neuropsychologia *48*, 2158–2166.
43. 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. Neurosci. *28*, 251–257.
44. 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.
45. Bottari, D., Troje, N.F., Ley, P., Hense, M., Kekunnaya, R., and Röder, B. (2016). Sight restoration after congenital blindness does not reinstate alpha oscillatory activity in humans. Sci. Rep. *6*, 24683.
46. Badde, S., Ley, P., Rajendran, S.S., Shareef, I., Kekunnaya, R., and Röder, B. (2020). Sensory experience during early sensitive periods shapes cross-modal temporal biases. eLife *9*, e61238.
47. Putzar, L., Gondan, M., and Röder, B. (2012). Basic multisensory functions can be acquired after congenital visual pattern deprivation in humans. Dev. Neuropsychol. *37*, 697–711.
48. Held, R., Ostrovsky, Y., de Gelder, B., Gandhi, T., Ganesh, S., Mathur, U., and Sinha, P. (2011). The newly sighted fail to match seen with felt. Nat. Neurosci. *14*, 551–553.
49. Chen, J., Wu, E.D., Chen, X., Zhu, L.H., Li, X., Thorn, F., Ostrovsky, Y., and Qu, J. (2016). Rapid integration of tactile and visual information by a newly sighted child. Curr. Biol. *26*, 1069–1074.
50. Papageorgiou, E., McLean, R.J., and Gottlob, I. (2014). Nystagmus in childhood. Pediatr. Neonatol. *55*, 341–351.
51. Watson, A.B., and Ahumada, A.J., Jr. (2005). A standard model for foveal detection of spatial contrast. J. Vis. *5*, 717–740.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Dataset with all experimental results	Mendeley Data	https://doi.org/10.17632/swrwmfy6tg.1
Software and algorithms		
MATLAB R2015b	MathWorks	https://www.mathworks.com/
R 3.3.3	R project	https://www.r-project.org/
RStudio 1.1.423	RStudio	https://www.rstudio.com/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed and will be fulfilled by the Lead Contact, Irene Senna (irene.senna@uni-ulm.de).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Full datasets including all the experimental results of the two experiments have been deposited on Mendeley at https://doi.org/10.17632/swrwmfy6tg.1. This study did not generate unique codes. Any additional information required to reanalyze the data reported in this paper is available from the Lead Contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Participants

Sixteen Ethiopian cataract-treated children and adolescents (mean age: 13.1, age range: 8–20 years, 1 left-handed, mean time since surgery: 2.2 years, range 3 days–10 years) took part in Experiment 1 (see Table S1 for details). We also sought to test 8 participants 2 days prior to their operation. Among these participants, 4 could not see the stimuli and were unable to perform the task prior to surgery. Results were therefore obtained from only 4 participants prior to surgery (mean age: 10.5, age range 8–15 years, all right-handed). Another subset of 7 participants was re-tested a second time 4 months after the first test (mean age: 11.9, age range 8–15 years, 1 left-handed, mean time since surgery for first test: 26 days, range 3 days–2 months). Twenty-three Ethiopian cataract-treated children and adolescents (age: 12, age range: 8–19 years, 1 left-handed, mean time since surgery: 1.93 years, range 1 day–10 years) took part in Experiment 2 (Table S1). Among them, 10 participants also took part in Experiment 1. A subgroup of 5 cataract-treated individuals was tested multiple times in follow-up tests of Experiment 2 taking place a few days (mean: 2.4 days, 1–3 days), 4 months, and 1 year after surgery.

An ophthalmologist assessed the visual abilities of the individuals with cataracts immediately before and after surgery (as soon as bandages were removed). Before surgery, he evaluated light perception, awareness of hand motion, and finger counting at different distances. All participants had at least light perception (for further details, see Table S1). They presented dense bilateral cataracts, which were classified as congenital, as suggested by the fact that each participant showed optical nystagmus, a signature of early-onset visual deprivation.[50] A cataract is classified as congenital when it is either present at birth or develops within the first months of life. Most participants had misaligned eyes (strabismus), and some had a family history of bilateral cataracts. Participants received ophthalmological examination, underwent cataract removal surgery, and received an intraocular lens implant at the Hawassa Referral Hospital several years after birth. We additionally tested spatial visual acuity by determining the cut-off frequency of the contrast sensitivity function (CSF) in all cataract-treated participants immediately before every experimental session and in a subset of participants immediately before surgery (see 'Method details, Assessment of the spatial visual acuity'). We did not measure the CSF in all pre-operated participants either because their vision was too poor or because the testing procedure was not available at the time of the surgery. In Experiment 1, the cataract-treated group showed a mean CSF cut-off frequency of 4.18 cpd (range: 1–13.9 cpd). The pre-operated participants taking part in Experiment 1 had a mean CSF of 1.33 cpd (range: 0.08–3.4 cpd). In the subset of 7 participants tested twice, the average visual acuity tended to improve slightly between the first and the second test four months later, from 3.3 cpd (range: 1.3–6 cpd) to 4.9 cpd (range: 1.35–6.2 cpd), although this trend did not reach significance (Wilcoxon signed rank test, $p = 0.13$). We compared the performance of the cataract-treated participants in Experiment 1 with that of a control group of sighted individuals in a comparable age range. Ninety-seven German (age: 12.1, 4–20 years, 12 left–handed) and 15 Ethiopian (age: 11.3, 8–15 years, all right-handed) sighted individuals participated in the study as a control group. Since the group of German sighted individuals did not differ from that of the Ethiopian sighted participants in any variable of interest, the two groups were merged. To assess whether any possible difference in performance between the cataract-treated and the sighted participants might merely be the result of overall lower visual acuity in the cataract-treated individuals, we tested a second control group of 16 German sighted participants (mean age: 13.3, 8–20 years, 1 left-handed) with matched visual acuity (mean: 4.3, 1.7–11 cpd). We blurred vision (see 'Method details, Experiment 1') in the sighted participants of this second control group to mimic the visual acuity of the cataract-treated participants. We individually matched the individuals of the sighted control group with the cataract-treated participants according to visual acuity and age.

In Experiment 2 participants presented a mean CSF cut-off frequency of 3.73 cpd (range: 0.039–12.1 cpd). The visual acuity of the subset of 5 participants tested three times did not show major changes over time, remaining stable across the first (4.24 cpd, 1.2–7.6), second (6.31 cpd, 1.5–12.7), and third (4.65 cpd, 0.8–8.7) tests. To further monitor the correlation between development of vision in terms of CFS and the visual JND_V measure in Experiment 2, 14 cataract-treated participants were re-tested 2, 3 or 4 times in another set of follow-up experiments (Figure S1E). Ethiopian participants performed the experiments in the Hawassa Referral Hospital, in the Shashamane Catholic School for the blind, or in the Sebeta Blind School. German participants were tested in primary and secondary schools in the southwest of Germany (Baden Württemberg and Bavaria) and at Ulm University. The procedure was approved by the ethics committee of the University of Bielefeld and Hawassa University. Participants' parents or legal guardians gave their written consent to the surgical treatment and the behavioral tests.

METHOD DETAILS

Experiment 1

Participants sat at a table with their head comfortably resting on a chin-rest. They faced the experimental setup, which consisted of 4 pillars (11 cm high) holding a 14 by 14 cm panel with an aperture (7x5 cm height x width) in its center for holding a lens (Figure 1A). Through the aperture they could see a rectangular object, placed centered on a stand right below the panel. A transparent Plexiglass optical lens was placed on top of the aperture at 20 cm distance from the participants' eyes. The lens could either be flat and thus not introducing any distortion to the visual image, or it could be half-cylindrical, thereby magnifying the visually observed object along its vertical axis with a magnification factor of roughly 1.75 (depending on the distance of the eye to the lens). In a match-to-sample procedure, participants were required to explore the standard object placed below the lens visually, haptically, or visual-haptically. They had to compare its size (i.e., its height) to that of a set of ten comparison objects. Standard and comparison objects were rectangular and made from hard plastic material. They were fixed with a small centered pole either to the setup holding the lens (standard object) or to the stand (comparison objects). To maximize the contrast between the objects and the surroundings, the stimuli were black, while the rest of the setup was white. All objects had a width of 4 cm but differed in height: the standard object was 2, 3, or 4 cm in height during training and 2 cm high during test (see below), while the comparison stimuli could have one of 10 heights linearly spaced between 1.2 and 4.8 cm. The comparison stimuli were aligned on the stand in ascending order. Participants were asked to close their eyes for the entire duration of the experiment and to open them only upon the experimenter's request. The experimenter positioned the standard object below the lens after covering the setup with a piece of black cloth. In the visual condition, the experimenter removed the cloth and asked participants to open their eyes and to observe the object, while keeping the head still on the chinrest. After around 5 s (as in Rock and Victor[26]), in which participants could visually explore the standard object, the experimenter covered the setup again with the piece of cloth. In the haptic condition, participants were required to use their dominant hand to grasp the object along its vertical axis with a precision grip (i.e., with index and thumb). The object could be easily reached through large openings on each side of the experimental setup. The experimenter made sure that participants were correctly grasping the object (i.e., along the requested axis and with the proper grip), and after around 5 s she would remove the participant's hand from the object. The setup was covered with the piece of cloth for the entire duration of each haptic trial in order to prevent participants from seeing the object in case they would accidentally open their eyes. In the visual-haptic condition, participants were instructed to grasp the object with a precision grip while watching. The experimenter removed the cloth from the setup as soon as the hand had reached the object, making sure that the object would be explored with vision and touch simultaneously. Participants were instructed to neither look away nor remove their fingers from the object during exploration. After around 5 s, the experimenter simultaneously covered the setup again and removed the participant's hand from the object. As soon as the participant had explored the standard object, the experimenter placed the stand holding the comparison stimuli in front of the participant, at the distance of the standard object. For the rest of the experimental session, the stand holding the comparison objects was hidden from view and out of participant's reach. For trials in which participants had to give a visual response, participants were instructed to carefully observe each object visually and to point toward the object that in their opinion matched the standard. For trials requiring a haptic response, the experimenter guided the participant's hand to the first comparison object (i.e., the smallest one). Participants kept their eyes closed and were asked to move their hand along the set of stimuli from the smallest to the biggest object, exploring each comparison object with a precision grip. They then select the one that in their opinion matched the standard object in size.

The experiment started with a short training session to help participants internalize the instructions and familiarize themselves with the task. Training was performed with the flat lens, which did not distort the object's apparent visual size. During training the standard object was explored either visually (V) or haptically (H) and had to be matched to the comparison stimulus either visually or haptically.

Four possible combinations were therefore possible (V-V: vision to vision; H-H: haptic to haptic; V-H: vision to haptic; H-V: haptic to vision). During training, one trial for each combination was presented. Either the V-V or the H-H trial was presented first, given that those two conditions were easier to comprehend for the participants. The order of the training trials was counterbalanced across participants. For each participant, two standard objects were presented (chosen randomly across the 3 standard stimuli), with each object being shown twice. If 4 trials were not enough for the participant to understand the task, another block of 4 trials was run.

After the training session, the experimenter removed the flat lens from the setup unbeknownst to the participant and introduced the distorting lens, which was used for the experimental session. In the experimental session, participants explored the standard object either visually, haptically, or visual-haptically, and they had to match this object to the set of comparison stimuli presented either haptically or visually. Four conditions (two unisensory and two multisensory) were tested: H-H (both standard and comparison objects were presented haptically); V-V (both standard and comparison objects were presented visually); VH-V (the standard object was presented visual-haptically, the comparison objects visually); VH-H (the standard object was presented visual-haptic, the comparison object haptically). The experimental session consisted of 3 blocks of 4 trials each, yielding a total of 12 trials. In each block, one trial of each of these four conditions was presented in random order. Unlike during training, we used only one standard object for testing (2 cm high) in order to keep the trial number low. After each trial, the experimenter removed the object from the setup and put it back in place to give participants the impression that the to-be-explored object would differ across trials.

To mimic the visual acuity shown by cataract-treated participants, we blurred vision in sighted controls. Blurring was achieved by placing several layers of semi-transparent foils under the optical lens (more layers generated more blur). In a pilot study using a group of sighted individuals (with normal or corrected to normal vision), we tested the number of layers needed to obtain each target CSF cut-off frequency value (i.e., each CSF cut-off value in the group of cataract-treated individuals, *cf.* 'Assessment of the spatial visual acuity'), and an analogous shape of the contrast sensitivity function. The CSF was measured after placing the blurring foils in front of the computer screen. The distance between the foils and the screen was the same as between the foils and the stimuli in the multisensory integration task. To ensure that this procedure would lead to the desired reduction in visual acuity, we then measured the reduced CSF in each sighted control and individually matched it for age and visual acuity to one of the cataract-treated participants.

Experiment 2

In a size discrimination task, participants compared the size of a standard stimulus (a 60 mm cube) with that of a set of comparison stimuli. In a 2-alternative forced-choice task, participants had to indicate which of the standard or comparison objects was larger. The stimuli consisted of a series of white 3D-printed cubes, which were presented haptically, visually, or visual-haptically. In contrast with Experiment 1, no discrepancy was introduced between vision and touch (i.e., the stimuli were not observed through a distortion lens).

For the first 12 of the 23 participants, we adopted an adaptive procedure. The experiment consisted of 3 blocks, one for each condition (visual, haptic, and visual-haptic), whose order was randomly varied across participants. Participants sat at a table in front of the experimenter, and the objects were placed within reaching distance of the participants. To maximize the contrast between the objects and the surroundings, the stimuli were presented against a dark background. The standard object (60 mm) was compared to a comparison cube, and afterward the participant had to indicate which of the two cubes was larger. The sizes of the comparison cubes ranged from 50 to 70 mm, in steps of 1.0 mm in the center region between 55 mm and 65 mm, followed by a 2 mm step (53 and 67 mm), and a 3 mm step (50 and 70 mm). The size of the comparison stimulus was changed between trials according to a 1-up/1-down staircase procedure converging on the point of subjective equality (PSE). Since this staircase method focused more on the central region around the PSE, while the Just Noticeable Difference JND (central to Experiment 2) is estimated less reliably, we slightly adapted the procedure for the following 11 participants to the more efficient method of constant stimuli for estimating the JND. The adaptive staircase procedure was useful in determining the exact range of stimuli to be used in the further tests.

The remaining 11 of the 23 participants took part in a version of the discrimination task in which we used the method of constant stimuli to estimate discrimination thresholds. The experimental procedure was otherwise similar: in the 3 blocks (visual, haptic, and visual-haptic) participants explored the standard (60 mm) and the comparison stimulus (53, 57, 59, 61, 63, or 67 mm) in random order and indicated which of the two objects was larger. At the beginning of each trial, the experimenter placed one object on the table in front of the participant within reaching distance (30-40 cm). The participant explored the object for around 5 s. The experimenter then removed the cube and positioned the second stimulus in the same location as the first and let the participant explore it for the same amount of time. After removal of the second object, the participant reported which of the two cubes was perceived to be larger. In the visual condition the participant was instructed to visually explore each cube without touching it. In the haptic condition, a black shield was positioned in front of the object to occlude it from sight. The participant reached for the cube from the side with their dominant hand. In the visual-haptic condition participants were instructed to haptically and visually explore the object simultaneously without looking away or removing their fingers from the object during the exploration. The experimenter took care to remove the shield at the same time that the participant's hand had reached the stimulus. Each standard-comparison pair was presented in randomized order 10 times for each of the 3 conditions, yielding a total of 180 trials. The experiment was conducted in 2 sessions, lasting around 30 min each.

For the analysis, the sets of JND data gathered from the two slightly different experimental procedures were joined, since they did not differ significantly from one another in any condition according to Wilcoxon rank-sum tests and they showed overlapping confidence intervals.

Assessment of the spatial visual acuity

We tested spatial visual acuity by determining the cut-off frequency of the contrast sensitivity function (CSF) with the procedure described in McKyton et al.[31] The participant's head was supported by a chin-rest at 30 cm distance from a gamma-corrected, 15.6-inch screen (1920 × 1080 pixels resolution). The distance was reduced to 15-20 cm for few participants with extremely low vision, who would have not been able to perform the task otherwise. Participants observed a sequence of Gabor patches, 19.5 cm diameter sinusoidal gratings with a Gaussian envelope, which were rendered at different contrast levels and with different spatial frequencies oriented either vertically or horizontally. In each trial, the participants had to report the orientation of the grating. In the initial part of the test, gratings were shown at 100% contrast. The grating tested in the first trial was at the lowest spatial frequency: 512 pixels per cycle, corresponding to 0.042 cycles per degree (cpd) at a viewing distance of 30 cm. If participants reported the correct orientation, the grating with the next higher spatial frequency was shown. Gratings could have one of 9 spatial frequencies spaced evenly on a logarithmic scale with the maximum of 2 pixels per cycle (10.75 cpd at 30 cm distance). Upon the first mistake, a 3:1 staircase procedure was introduced. Three consecutive correct responses led to the next higher frequency, while one mistake led to the next lower frequency. The procedure stopped after 6 reversals. In the second part of the CSF test, we determined the detection threshold for each frequency in a different block; that is, we kept the frequency of the gratings constant, while varying the contrast. We tested the frequencies one after the other from lowest to highest. For each spatial frequency, the first patch was shown at 100% contrast. A total of 8 contrast levels were used, spaced logarithmically between 100% and 0.78%. As long as the participant's response was correct, the contrast was progressively reduced by one level after each response. Upon the first error, a 3:1 staircase procedure was introduced, similar to the one described above, to measure participant's threshold contrast for each frequency. If no error occurred, the procedure was stopped at a minimum contrast of 0.78%. We computed the CSF for each participant by converting the sensitivity (1/contrast at threshold) in a logarithmic scale for each frequency and plotted it as a function of the spatial frequency. We then fitted the CSF for each participant with an inverse parabola[31,51] to get the CSF cutoff frequency, i.e., the highest spatial frequency that the participant could still see with the maximal contrast. The test was performed binocularly, as it happened for the main experiments.

QUANTIFICATION AND STATISTICAL ANALYSIS

Experiment 1

In each trial, we recorded the participant's response as the height of the selected comparison object. The physical standard object was 2 cm high, and it visually appeared as being roughly 3.5 cm high (due to the lens' magnification factor = 1.75, thus leading to 2*1.75 = 3.5 cm). As the magnification factor depends on the distance between the lens and the eye and the head was only fixed by a chin rest, there was some little possibility for head movements slightly affecting the magnification. Therefore, the mean of the size estimates in the V-V condition slightly deviates from 3.5cm. The mean in each of the 4 conditions was calculated for each participant. A Friedman test for intra-group comparison was conducted on the participants' responses with Condition as factor. Significance was set at $p = 0.05$. We accounted for multiple comparisons here and throughout the study by using Bonferroni-Holm corrected Wilcoxon signed-rank tests.

To compare the performance across the three groups (cataract-treated, sighted controls tested under normal visual conditions, and sighted controls tested with visual blur) we calculated two measures: the average visual weight across the two multisensory conditions ($\overline{w}_V$) and the multisensory influence (*MI*).

The first measure $\overline{w}_V$ indicates how much vision influences touch when integrating the multisensory signals. We computed the average visual weight as:

$$\overline{w}_V = \frac{\widehat{S}_{VH} - \widehat{S}_H}{\widehat{S}_V - \widehat{S}_H} \quad \text{(Equation 2)}$$

where $\widehat{S}_{VH}$ is the average size estimate across the two multisensory conditions VH-H and VH-V, $\widehat{S}_H$ is the size estimate from the unisensory haptic condition H-H, and $\widehat{S}_V$ is the size estimate from the unisensory visual condition V-V. If in the multisensory conditions a participant relied on vision only ($\widehat{S}_{VH} = \widehat{S}_V$), the average visual weight is $\overline{w}_V = 1$. If only touch is used for the multisensory estimates ($\widehat{S}_{VH} = \widehat{S}_H$), $\overline{w}_V = 0$. If both senses are used, they are weighted according to: $\widehat{S}_{VH} = \overline{w}_V \widehat{S}_V + \overline{w}_H \widehat{S}_H$, with the visual and haptic weights adding to 1: $\overline{w}_V + \overline{w}_H = 1$.

We calculated $\overline{w}_V$ for each participant and compared it across the three groups via a Kruskal–Wallis test, followed by Bonferroni-Holm corrected Wilcoxon rank-sum test.

The second measure is the Multisensory Influence (*MI*) of vision and touch on the combined estimate. It was calculated as:

$$MI = w_V(VH\text{–}H) + w_H(VH\text{–}V) \quad \text{(Equation 3)}$$

where $w_V(VH\text{–}H)$ is the visual weight in the VH-H condition (visual influence on touch), and $w_H(VH\text{–}V)$ is the haptic weight in the VH-V condition (haptic influence on vision). These weights were calculated as:

$$w_V(VH\text{–}H) = \frac{\hat{S}_{VH-H} - \hat{S}_H}{\hat{S}_V - \hat{S}_H} \quad \text{(Equation 4)}$$

and

$$w_H(VH\text{–}V) = \frac{\hat{S}_{VH-V} - \hat{S}_V}{\hat{S}_H - \hat{S}_V} \quad \text{(Equation 5)}$$

where $\hat{S}_{VH-H}$ and $\hat{S}_{VH-V}$ are size estimates in the VH-H and VH-V conditions, respectively. *MI* ranges between 0 and 1: If participants do not integrate vision and touch, i.e., there is no influence of vision on touch or vice versa, $\hat{S}_{VH-V} = \hat{S}_V$ and $\hat{S}_{VH-H} = \hat{S}_H$, therefore w_H(VH-V) = 0 and w_V(VH-H) = 0, leading to *MI* = w_V(VH-H)+w_H(VH-V) = 0. Instead, if participants mandatorily fuse the visual and haptic inputs, i.e., vision and touch influence each other to a degree that both visual and haptic estimates are identical irrespective of the magnification added to the visual input, *MI* = w_V(VH-H)+w_H(VH-V) = 1. Values slightly below 0 or above 1 (*cf.* Figure 1E) are attributed to noise.

We calculated *MI* for each participant and compared it across the three groups via a Kruskal-Wallis test, followed by Bonferroni-Holm corrected Wilcoxon rank-sum test.

In the cataract-treated group, we ran a multiple regression analysis (with Bonferroni-Holm corrected p values) on *MI* with age at test, time since surgery, and visual acuity (i.e., log-transformed CSF cut-off frequency) as factors. These different factors were not significantly correlated with each other (visual acuity versus time since surgery, r = -.12, p = 0.67; visual acuity versus age, r = -.07, p = 0.79; age versus time since surgery, r = 0.45, p = 0.09). Note that considering age at surgery instead of age at test in the multiple regression revealed a similar trend (t = 2.13, p = 0.057, *cf.* main text) as both measures are highly correlated (.59, p = 0.02). One participant (p17, see Table S1) was not included in this multiple regression, since we did not have all the demographic information about him necessary for this analysis. Results were analyzed using MATLAB (Mathworks, Natick, MA).

Experiment 2

According to the "maximum-likelihood estimation" (MLE) model, when inputs from different senses *i* are integrated optimally, the minimum variance unbiased estimate $\hat{S}$ is a weighted average of the individual sensory inputs $\hat{S}_i$, with weights w_i proportional to their reliability (inverse variance).[1] Each estimate is inherently noisy, and if the noises of the sensory estimates $N_i(0, \sigma_i^2)$ are unbiased, normally distributed, and independent, then the maximum-likelihood estimate is given by:

$$\hat{S} = \sum_i w_i \hat{S}_i \quad \text{(Equation 6)}$$

with

$$w_i = \frac{1/\sigma_i^2}{\sum_j 1/\sigma_j^2} \quad \text{(Equation 7)}$$

When combining unisensory estimates (e.g., visual $\hat{S}_V$ and haptic $\hat{S}_H$), the resulting multisensory estimate (visual-haptic $\hat{S}_{VH}$) is more precise (i.e., has a smaller variance) than each unisensory estimate, with variance:

$$\sigma_{VH}^2 = \frac{\sigma_V^2 \sigma_H^2}{\sigma_V^2 + \sigma_H^2} \quad \text{(Equation 8)}$$

We fitted cumulative Gaussian distributions to the proportion of "comparison bigger" responses for each condition and participant.[37] To explore whether the multisensory presentation led to a more precise estimate compared to when using a single sense alone, as it would be expected according to MLE rules, we determined size discrimination thresholds (i.e., the Just Noticeable Differences, JNDs) in each of the three conditions (vision, haptic, and multisensory). JNDs for each participant and condition were calculated from the psychometric curves by halving the difference between the comparison size at which participants made 16% and 84% "comparison bigger" responses. For fitting, the Point of Subjective Equality (PSE) could be fixed to the size of the standard at 60 mm, so with the JND there was only one free parameter. We also predicted the multisensory visual-haptic JND from the MLE rules as follows:

$$JND_{optimal} = \sqrt{\frac{JND_V^2 JND_H^2}{JND_V^2 + JND_H^2}} \quad \text{(Equation 9)}$$

Equation 9 is derived from Equation 8 knowing that for 2-interval forced-choice tasks during which 2 noisy estimates are compared, the JND calculated between the 0.16 and 0.84 points (16% and 84% "comparison bigger" responses) corresponds to: $JND = \sqrt{2}\sigma$ (cf. Ernst[32] for further details).

Participants' JNDs in the three conditions were compared via Friedman's test, followed by Bonferroni- corrected Wilcoxon signed-rank tests.

In a further step, we investigated whether visual acuity, age, and time since surgery could affect participants' ability to optimally integrate multisensory signals. In this case, the multisensory estimate should be better than either of the unisensory estimates. We calculated the integration gain $g_{integration}$ as:

$$g_{integration} = log\frac{min(JND_V, JND_H)}{JND_{VH}}, \quad \text{(Equation 10)}$$

where any advantage of the multisensory condition over the better unisensory JND would lead to positive values and would indicate integration, while a higher JND_{VH} in the multisensory condition would lead to negative values, which would be a sign of interference. We also investigated the possible contribution of the factors of age, time since surgery, and visual acuity on the integration gain by fitting an exponential function separately for each of the three factors, as in Equation 1 (main text), where *c* was fixed to the predicted optimal integration gain $\overline{g}_{optimal}$ of the group. The predicted optimal gain of the group was calculated for each participant as the log-transformed ratio between the better (i.e., lower) unisensory JND and the visual-haptic JND predicted by MLE (i.e, $JND_{optimal}$ rather than the empirical JND_{VH}), and then averaged across participants (group optimal predicted integration gain, $\overline{g}_{optimal} = 0.066$). We also calculated the difference between the predicted and the observed gain $\Delta g_{integration}$ in each participant (thus, $\Delta g_{integration} = 0$ signifies optimal integration of the multisensory signals). We fitted an exponential between $\Delta g_{integration}$ and the time since surgery using: $\Delta g_{integration} = ae^{-bx}$. The difference between the predicted and observed integration gain tended to exponentially decrease with time since surgery and approached 0 (i.e., optimal performance, learning rate $b = 0.3$, 95% confidence interval, CI = [-.39, 0.98], Figure 2B, inset). Results were analyzed using MATLAB (Mathworks, Natick, MA). Psychometric functions were fitted using R (http://www.r-project.org).