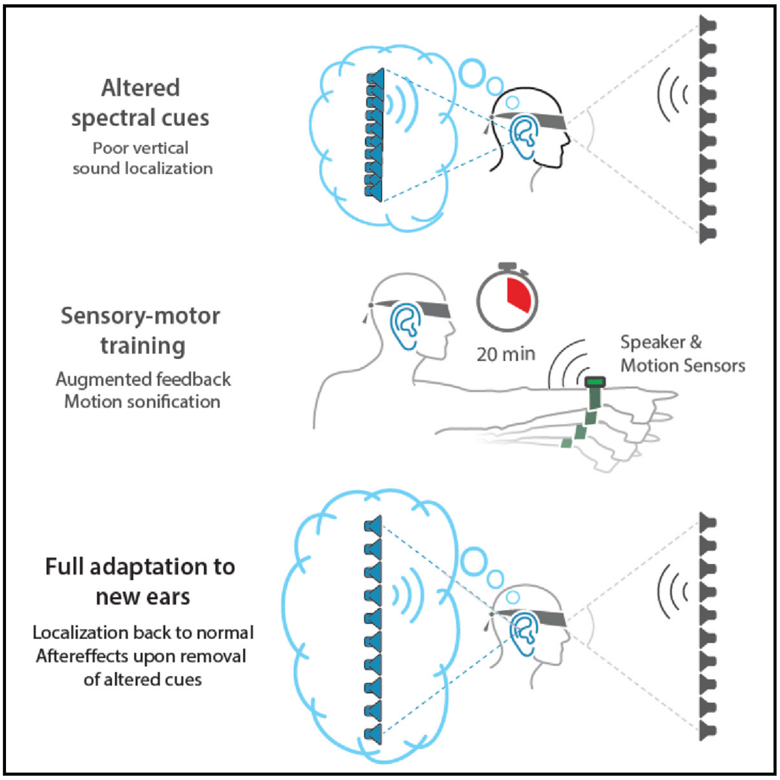
iScience

Article

Happy new ears: Rapid adaptation to novel spectral cues in vertical sound localization

Graphical abstract



Highlights

- Disruption of HRTFs and vertical auditory localization after wearing new-ears
- Sensorimotor-training induced adaptation to the new-ears for vertical localization
- Aftereffects after new-ears removal resulting in precision drop for few trials

Authors

Cesare Parise, Monica Gori, Sara Finocchietti, Marc Ernst, Davide Esposito, Alessia Tonelli

Correspondence

monica.gori@iit.it

In brief

Neuroscience; Bioengineering

Parise et al., 2024, iScience *27*, 111308
December 20, 2024
https://doi.org/10.1016/j.isci.2024.111308

CellPress

iScience

CellPress
OPEN ACCESS

Article

Happy new ears: Rapid adaptation to novel spectral cues in vertical sound localization

Cesare Parise,[1,5] Monica Gori,[2,5,6,*] Sara Finocchietti,[2] Marc Ernst,[3] Davide Esposito,[2] and Alessia Tonelli[2,4]
[1]Department of Psychology, University of Liverpool, Liverpool, UK
[2]Unit for Visually Impaired People, Italian Institute of Technology, Genoa, Italy
[3]Department of Psychology, University of Ulm, Ulm, Germany
[4]School of Psychology, University of Sydney, Sydney, Australia
[5]These authors contributed equally
[6]Lead contact
*Correspondence: monica.gori@iit.it
https://doi.org/10.1016/j.isci.2024.111308

SUMMARY

Humans can adapt to changes in the acoustic properties of the head and exploit the resulting novel spectral cues for sound source localization. However, the adaptation rate varies across studies and is not associated with the aftereffects commonly found after adaptation in other sensory domains. To investigate the adaptation' rate and measure potential aftereffects, our participants wore new-ears to alter the spectral cues for sound localization and underwent sensorimotor training to induce rapid adaptation. Within 20 min, our sensorimotor-training induced full adaptation to the new-ears, as demonstrated by changes in various performance indexes, including the localization gain, bias, and precision. Once the new ears were removed, participants displayed systematic aftereffects, evident as drop in the precision of localization lasting only a few trials. These results highlight the short-term plasticity of human spatial hearing, which is capable to quickly adapt to spectral perturbations and inducing large, yet short lived, aftereffects.

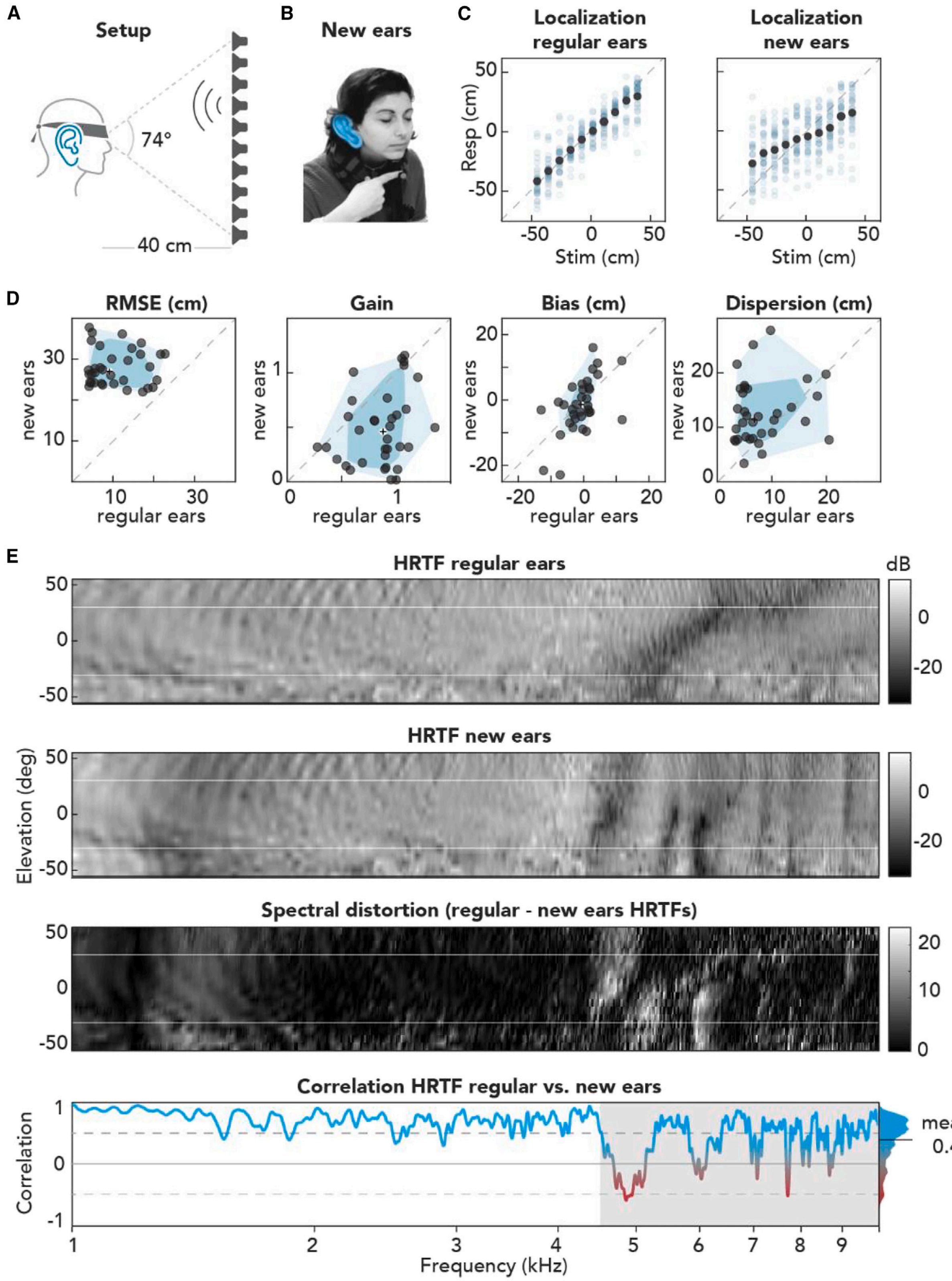
INTRODUCTION

The plasticity of perceptual systems is not limited to the early stages of life[1]; it is a lifelong process[2,3] that affects all perceptual domains. This includes spatial hearing, i.e., the ability to localize sounds and create a mental map of the sound sources in space. Such abilities rely upon spectrotemporal cues emerging from the acoustic properties of our pinnae and body, known as the head-related-transfer functions (HRTFs). Several studies have investigated how manipulations of the HRTFs affect spatial hearing and how participants adapt to such changes by learning new maps.[4,5] The experimental manipulations of the HRTFs range from using hearing aids or plugs[6,7] to long-term or intermittent monaural blocks[8] or modification of the pinnae with semi-permanent molds.[9–12] Results demonstrate that participants can adapt to all such manipulations, though the rate of adaptation varied dramatically across studies. For instance, Florentine[13] applied a monaural earplug for a period of 27–101 days and saw that after 4–10 days, participants presented a partial adaptation to the auditory image. Hofman et al.[12] used molds designed to manipulate the HRTFs and found that sound localization along the sagittal plane was compromised right after fitting the mold, while localization along the horizontal plane was relatively unaffected. Without training, over time (range 23–39 days), participants learned to use the novel HRTFs.

Recent studies exploited closed-loop training protocols, where participants were actively engaged in sound localization tasks while receiving constant feedback on their performance to speed-up adaptation.[14] Irving and Moore[15] had participants wear a plug on one ear and measured sound localization on the horizontal plane. During training, participants received feedback on their performance (i.e., correct/incorrect) and quickly adapted to monaural deprivation, reaching a plateau after four days. That is 5–8 times faster than Hofman et al.[12]

Further studies explored the effects of feedback and multisensory training on adaptation to altered auditory spatial cues. While feedback on the accuracy of behavioral responses induced adaptation to monaural deprivation after 4 days, multisensory and sensorimotor training protocols seem to be the fastest ones.[16] For instance, recent VR studies demonstrate that significant adaptation to generic (i.e., non-personalized) HRTFs[17] or monaural listening[18,19] can occur within minutes of sensorimotor training with multisensory feedback. The use of generic HRTFs, however, does not represent a dramatic manipulation of an individual's spectral cues: generic HRTFs are successfully used in most VR applications. Hence, it is unclear whether the auditory system can quickly adapt to more drastic manipulations of the HRTFs.[12,13,20]

A unique aspect of adaptation to altered auditory spatial cues is the absence of measurable aftereffects. Indeed, all studies that assessed what happens to spatial hearing once the manipulations are removed failed to find any difference in performance with the baseline condition, i.e., the performance before the manipulation.[4] This finding was compared by



(legend on next page)

Hofman et al.[12] by drawing with learning a new language, where improved proficiency in the new language does not affect proficiency in the native one. While intriguing, more evidence is necessary to support such an interpretation, which contrasts with most instances of adaptation described in sensory neuroscience.[21–23] Auditory aftereffects are present for low-level stimulus properties[24,25] and high-level information.[26] Therefore, one might wonder whether the lack of after-effects induced by adaptation to novel spatial cues is a genuine phenomenon or whether they are present but undetected because the experimental or analytical tools were not sufficiently sensitive to measure them reliably.

To address these points, we conducted two experiments in the present work. In the first experiment, we tested a group of participants in a vertical auditory localization task before and after wearing a pair of new ears used to modify the shape of the pinna and, hence, the HRTFs (Figures 1B–1E). The aim was to quantify to what extent such manipulation could alter the localization performance. Next, in the second experiment, we took a subsample of participants to evaluate the effect of a quick audio-motor training protocol (lasting around 10 min) on the adaptation to the new HRTFs. The aim was to test whether such quick training sessions could prompt the spectral cues remapping and, at the same time, to test for the presence of aftereffects related to the exposure to the altered spectral cues with and without training. For the training, we used audio-motor inputs delivered by a wrist-worn audio device that played sounds only while moving.[27] This type of training has been proven effective for improving auditory localization skills in both typical individuals and blind people.[28–30]

We focus on vertical auditory localization for two reasons: firstly, elevation estimates rely on HRTFs; secondly, training and experimental sessions targeting one spatial dimension become quicker, thereby enabling the assessment of rapid adaptation and its potential aftereffects. We anticipate that our sensorimotor training led to a fast improvement in vertical auditory localization, and systematic aftereffects were present once the new ears were removed.

RESULTS

Experiment 1

To assess the impact of the new ears on the spectral cues for vertical sound localization, we measured the HRTF from a listener with and without modified ears. HRTFs were recorded in the same room where we conducted the experiment, a quiet testing space with no acoustic treatment, thus with the presence of background noise and reverberation of the room. Frequency sweeps were played from each loudspeaker while we recorded the sound from a reference microphone and binaural microphones placed inside the ear canal of a static listener (with and without modified ears). HRTFs were calculated for each source elevation as the ratio of the power spectra of the in-ear recordings to those of the reference recordings. This procedure was repeated separately for new and regular ears, and the resulting HRTFs (dB range: $-20,20$) and spectral distortion are represented in Figure 1E. To quantify the difference between regular and new HRTFs, we calculated the Pearson correlation between the two HRTFs for each frequency (Figure 1E, bottom). Results demonstrate that the new ears have a disruptive effect on the HRTFs above 4500Hz, where the average correlation was 0.4 and the minimum was -0.65. Notably, the new ears strongly affected the spectral structure of the HRTF: this can be appreciated from the variability in the correlation of the HRTFs across neighboring frequencies.

Concerning the behavioral assessment of the new ears-induced HRTF alteration on vertical localization, the response pattern (Figure 1C) shows that participants mis-localize sounds more with the new ears on than without them. Looking at the overall performance given by the *RMSE* (Figure 1D – see data analysis section for *RMSE*), we found that most participants have a low *RMSE*, i.e., good performance overall, in vertical sound localization. Nonetheless, with the new ears on, the *RMSE* increases significantly ($t_{33} = 13.15$, $p < 0.001$, d = 0.46, CI [-2.89 -1.61]), indicating a performance drop.

For a more detailed performance assessment, we calculated the linear fit between the perceived position and the real position of the sound for each participant, obtaining three main variables: *gain*, *bias*, and *dispersion* (see data analysis section for more details). In the session with new ears, the slope of the linear fit is flatter than in the session with regular ears, which represents a decrease in localization *gain* (Figure 1D - $t_{33} = 6.07$, $p < 0.001$, d = 1.04, CI [0.62 1.455]), while we found no difference in the *bias* at 0° (Figure 1D - $t_{33} = 0.6$, $p = 0.55$, $d = 0.1$, *CI* [-0.235 0.44]). The effect of ear manipulation is also present in the precision with which the participants performed the task, with larger *dispersion* for new as compared to regular ears (Figure 1D - $t_{33} = 3.9$ $p < 0.001$, $d = 0.67$, *CI* [-1.04 -0.29]).

The change in HRTFs resulting from wearing a new set of ears, as shown in Figure 1E and is reflected in a general reduction in vertical sound localization performance. Specifically, the perception of most participants appears to be more "noisy," with a decrease in precision (inverse of dispersion) and a gain reduction rather than a shift of the bias.

Figure 1. Setup and results of Experiment 1

(A) Representation of the setup.

(B) New ears used in Experiments 1 and 2.

(C) Localization results, with perceived position plotted against the physical position. The scatter plot on the left panel displays the performance with the regular ears, and the one on the right displays the performance with the new ears. Black circles represent the average among all participants, while gray circles are the points for each participant. The dotted line represents the equality line.

(D) Localization performance with regular and new ears, as assessed using the indexes *rmse*, *gain*, *bias*, and *dispersion,* displayed as bagplots.

(E) HRTFs of regular and modified ears, the spectral distortion, and the correlation between regular and new ears for each frequency. The marginal on the right represents the histogram (and mean) of the correlation in the frequencies above 4500Hz. The white lines in the HRTFs represent the range of elevations tested in our experiments. See also Figure S1 in the supplementary materials.

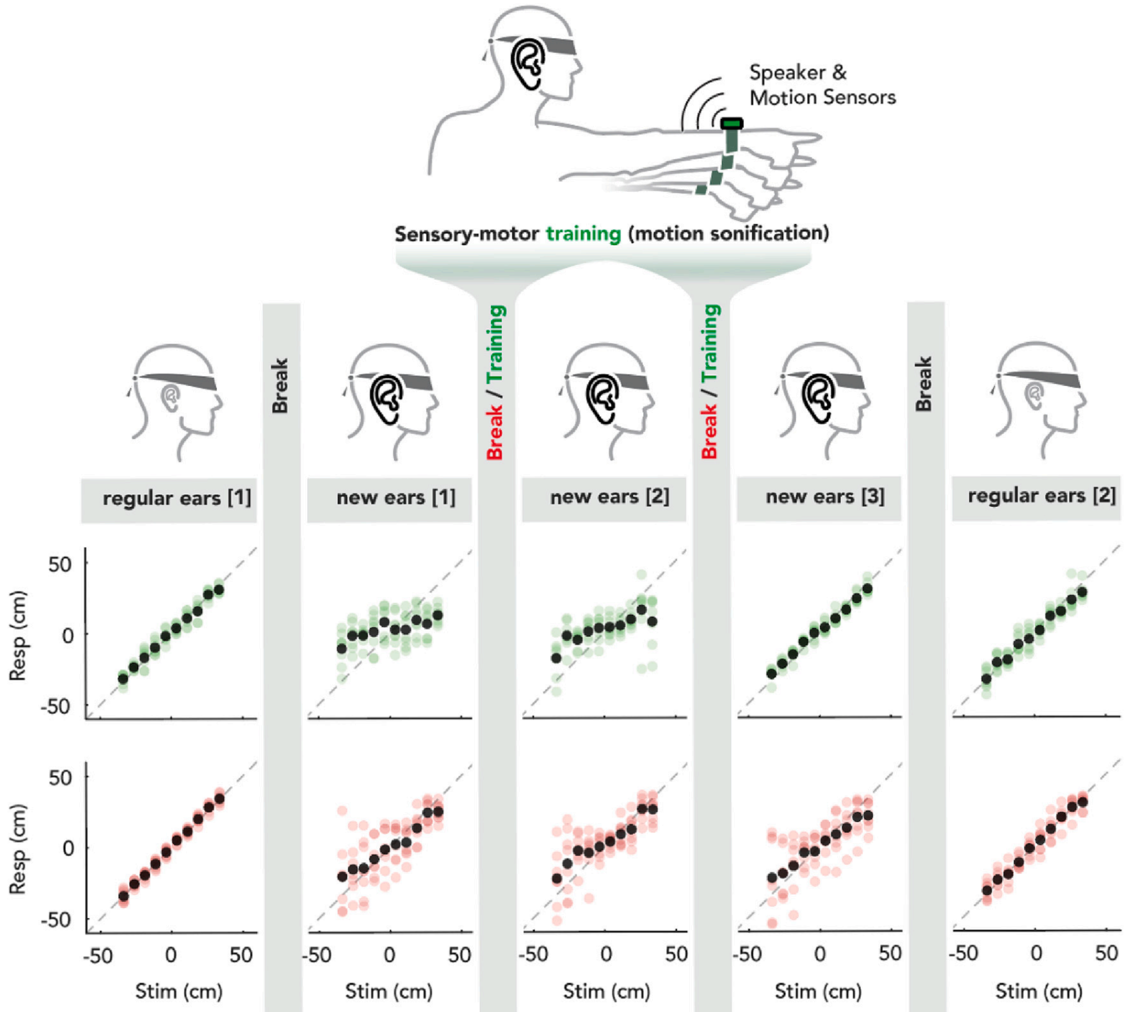


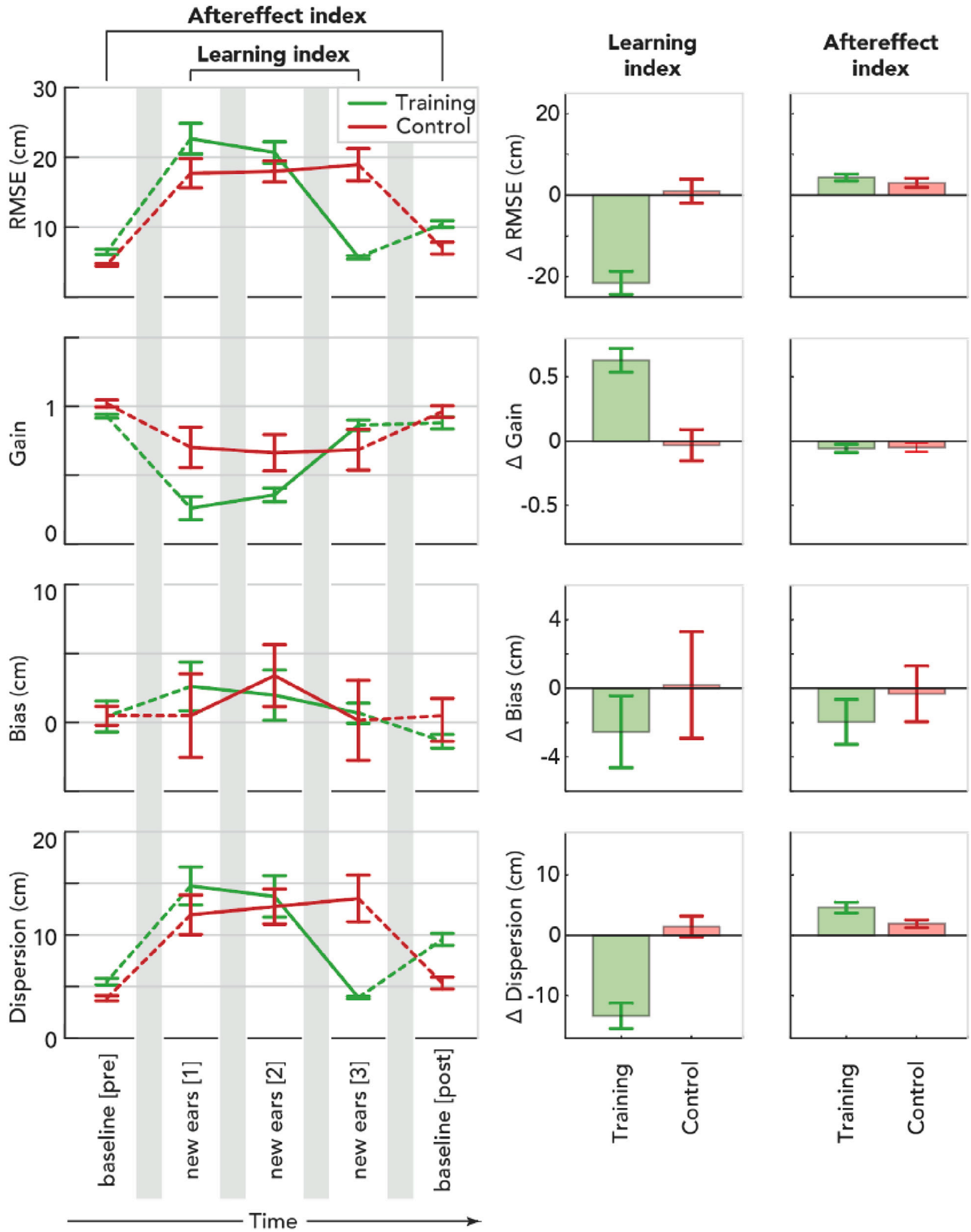
Figure 2. Results of Experiment 2

Top. Schematic representation of the training protocol using sonification of hand movements. Bottom. Localization performance in the five sessions of Experiment 2. Green points represent the responses of participants that performed the training protocol between each session, while the red dots represent the responses of the control group. Green and red dots represent the average response of each participant, while black dots are the global mean. See also Figures S2 and S3 in the supplementary materials for each participant.

Experiment 2

In experiment 2, we tested the effect of the audio-motor training sessions with new ears (spectral cues), as well as the effect of rapid adaptation and potential aftereffects. We selected a subsample of participants from experiment 1 and split it into training and control groups. Both groups underwent five vertical localization sessions with (central sessions, *new ears 1–3*) and without (first and last session, *regular ears 1*, *2*) new ears. The training group in-between central sessions underwent an audio-motor training wearing an auditory bracelet, while the control group performed the same activity without the bracelet. Figure 2 shows the experiment structure and an overall representation of the raw responses provided by the two groups in the vertical location sessions.

We use the same approach as experiment 1 to assess the performance within-session. For each session and participant, we calculate the *RMSE* and the best linear fit between the perceived position and the real position of the sound, obtaining the gain, bias, and dispersion (Figure 3 on the left). Furthermore, we calculated two new indices: *Learning index* and *Aftereffect index* (Figure 3 on the right). The *Learning index* is obtained by the difference of the given variable value between the second (*new ears 1*) and fourth session (*new ears 3*) and characterizes the effects of the training performed with the audio bracelet. The *Aftereffect index* is the difference for the given variable value between the first (*regular ears 1*) and last (*regular ears 2*) session and characterizes the effects of the prolonged exposure to the altered HRTF (for detailed analysis of all sessions, see Supplementary Materials). The significance of the between-group differences in the variables' *Learning indices* and *Aftereffect indices* were tested using unpaired t tests.



(legend on next page)

Regarding the *Learning index*, the *RMSE* (Figure 3) is significantly lower in the training group than in the control group (t_{14} = 5.67, p <0 .001, d = 0.87). The training group's negative *Learning index* indicates an improvement in the *RMSE* after the training compared to the control group, which is around zero, indicating no improvement. For the *gain*, the *Learning index* is significantly higher in the training group compared to the control group (t_{14} = 4.17, p <0 .001, d = 2.09), reflecting a training-induced reduction of the elevation compression. It should be noted that the *gain* in n*ew ears 1* was lower for the training group than the control group: while this implies reduced margins for improvement, it is nonetheless clear that the gain simply did not improve in the controls, while learning was nearly complete in the training group. The *Learning index* of the *dispersion* is significantly lower in the training group compared to the control group (t_{14} = 5, p <0 .001, d = 2.5), indicating a training-induced precision restoration. The variable that does not seem to be affected by training is bias, which remains constant for both groups (t_{14} = 0.21, p = 0.84, d = 0.104).

Next, we tested the effects of removing the new ears following the various training sessions (or lack thereof) by calculating the *Aftereffect index* (Figure 3). For the *dispersion* (Figure 3), we found that it was higher in the training group than in the control group (t_{14} = 2.72, p <0 .05, d = 1.36), indicating an aftereffect of training even after removing the new ears. Conversely, the *RMSE*, *gain,* and *bias* did not differ between the two groups (*RMSE*: t_{14} = 1.37, p = 0.19, d = 0.69; *gain*: t_{14} = 0.17, p = 0.865, d = 0.09; *Bias:* t_{14} = 0.85, p = 0.41, d = 0.425), meaning that the participants' performance has returned to the initial performance since this indexes are around zero.

As a final analysis, to assess the aftereffect duration, we compared our four dependent variables between the first and second halves of the last session (*Regular Ear 2*) across the two training conditions (Figure 4) using mixed ANOVAs. In line with previous results, we did not find a difference for factor *group* ($F_{1,14}$ = 1.64, p = 0.22, η^2 = 0.1), *block* (i.e., First vs. Second Half; $F_{1,14}$ = 1.47, p = 0.24, η^2 = 0.1) nor significant interaction ($F_{1,14}$ = 0.14, p = 0.72, η^2 = 0.01) for the *gain*, and also the *bias* (*group*: $F_{1,14}$ = 0.1, p = 0.75, η^2 = 0.007; *block*: $F_{1,14}$ = 4.34, p = 0.06, η^2 = 0.24; interaction: $F_{1,14}$ = 2.19 p = 0.16, η^2 = 0.13). Instead, for the *RMSE*, both main effects (*group*: $F_{1,14}$ = 8.54, p <0 .05, η^2 = 0.38; *block*: $F_{1,14}$ = 42.1, p <0 .001, η^2 = 0.75) and the interaction were significant ($F_{1,14}$ = 73.06, p <0 .001, η^2 = 0.84). Post-hoc analyses showed that the *RMSE* of the first block of the training group was significantly higher both compared to the second block (t_{14} = 10.63, p <0 .001, d = 3.9) and compared to the first block in the control group (t_{14} = 6.93, p <0 .001, d = 3.46), while there was no difference between the first and second block within the control group (t_{14} = 0.53, p = 1, d = 0.53, Figure 4, top). A similar result was also found for the *Dispersion* (Figure 4, bottom), with significant main effects (*group*: $F_{1,14}$ = 19.29, p <0 .001, η^2 = 0.58; *block*: $F_{1,14}$ = 28.93, p <0 .001, η^2 = 0.67) and the interaction ($F_{1,14}$ = 46.63, p <0 .001, η^2 = 0.77). The results were also confirmed in the post-hoc analysis, with the dispersion being significantly higher in the first block of the training group compared to both the second block of the same group (t_{14} = 8.63, p <0 .001, d = 4.06) and the first block of the control group (t_{14} = 7.82, p <0 .001, d = 3.91), while there was no difference between the first and second block of the control group (t_{14} = 0.29, p = 1, d = 0.15).

From these results on the Learning index, we can conclude that the audio-motor training appears effective in prompting the adaptation to the new HRTFs. However, as shown in Figure 3 (left column), it should be emphasized that this improvement is only evident following the second training (*new ears* [3]). In contrast, after the first training (new ears [2]), a performance improvement begins to be observed but is not significant (see supplementary materials for detailed analysis). This could stem from the initial adjustment period required for adapting to new HRTFs, during which information is sampled to construct a spatial map. Alternatively, the delayed improvement in behavior may result from participants requiring a certain level of error signals before the initiation of the adaptation to the new ears. Nevertheless, this interpretation of the results would need a follow-up study to investigate the nature of the effect.

From the results on the *Aftereffect index*, we can conclude that restoring normal ears reduces precision, but only for participants who underwent sensorimotor training. However, upon checking this effect based on the trials, we see that this effect vanishes within a few trials.

DISCUSSION

Using novel spectral cues, we investigated how humans learn to localize sounds on the vertical plane. To help participants adapt to the novel cues, we used sensorimotor training in which blindfolded participants wore a wristband emitting broadband noise when moving their arms. Within 20 min of training, participants fully adapted to the novel spectral cues and learned to localize sounds wearing the new ears as well as they did with their unmodified ears. Notably, upon removal of the new ears, we found systematic aftereffects in the form of reduced precision (increased dispersion) in sound localization with unmodified ears. Critically, adaptation and aftereffects only occurred in participants who underwent the auditory-proprioceptive training, demonstrating the systematic effects of perceptual learning in vertical sound localization.

Previous studies demonstrated that humans could learn to localize sounds with altered spectral cues using a variety of manipulations.[4] Some of them were relatively subtle, such as using generic instead of personalized HRTFs in virtual reality[10,31]; others were more severe and involved the insertion of molds in the pinnae to alter the geometry of the outer ears.[12,13] While in all cases, participants learned to use the novel spectral cues,

Figure 3. Localization performance across sessions as measured using *rmse*, *gain*, *bias*, and *dispersion*

Green lines represent the training group, while red lines represent the control group. The Learning Index is calculated as the difference in performance between the session "new ears 1" and "new ears 3". At the same time, the Aftereffect Index represents the performance difference between the sessions "baseline (pre)" and "baseline (post)". Error bars represent the standard error of the mean.

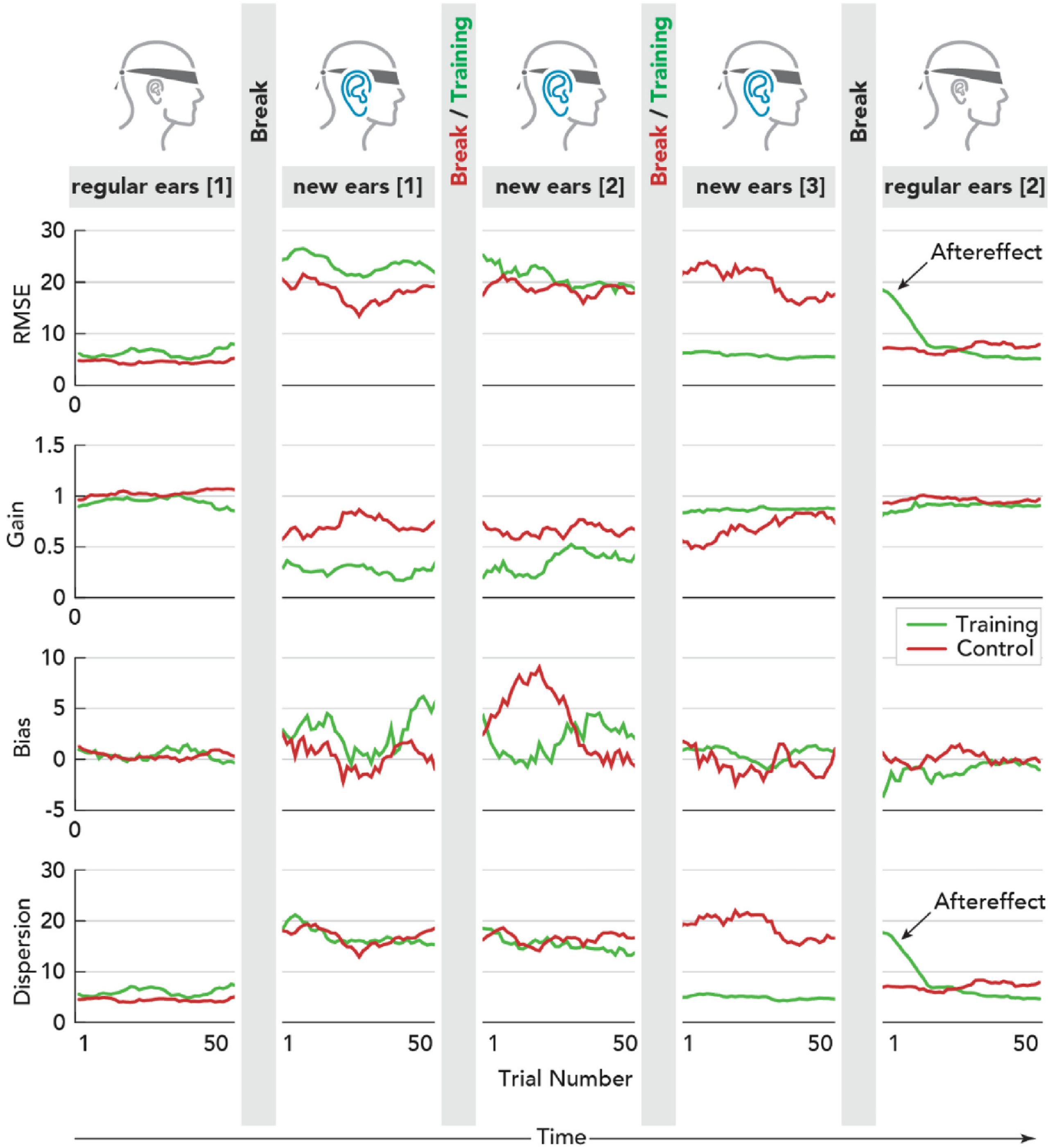


Figure 4. Build-up of learning effect over trials

The running average of the four summary statistics was obtained by pulling together data from all participants on a running window of 10 consecutive trials.

the speed of learning differed across studies, ranging from a few minutes to adapt to generic HRTFs in VR to several weeks in the case of more severe manipulations. Moreover, the learning was supported by visual cues, whose role in sound localization is well documented.[32–34] The present study extends previous results by demonstrating that humans can quickly adapt to dramatic manipulations of the HRTFs within minutes while solely relying on auditory-proprioceptive training without any visual cues.

Vertical sound localization relies on the spectral cues provided by the HRTFs. When the new ears alter such cues, participants immediately lose the ability to localize the sound in the vertical plane. This is because the new ears modify the natural mapping between spectral information and the relative location of the sound concerning the listener. The new mapping, however, can be re-learned with auditory-proprioceptive training: when participants move their hand with an auditory noise attached to the wrist, the sounds' spectra reaching the listeners' ears vary depending on the position of the wrist in the peri-personal space. Given that, participants can map near-field spectral cues to positions in space through proprioception, thereby quickly adapting to novel spectral cues. The present study only investigated the auditory near-field (both in the training and the experimental tasks), and it is a question for future research to test whether the same learning also transfers to the far-field.

Numerous studies have shown how vision influences hearing in various perceptual phenomena, such as sound localization[34,35] or auditory scene analysis.[36] However, multisensory research has only recently started addressing the relationship between the auditory and somatosensory systems.[37,38] It has been shown that the somatosensory system can influence audition not only from a behavioral point of view, such as in localization[39,40] but also at the cortical level, as bidirectional connections have been found between acoustic and motor regions in the brain.[41] Moreover, computational studies on the development of acoustic spatial representations have hypothesized that a stable representation of acoustic space in the brain can develop even without visual feedback but through non-supervised sensorimotor-learning based on dynamic acoustic inputs from the animal's own movements.[42,43] Our second experiment represents experimental evidence supporting this hypothesis: participants could reshape their spatial representation thanks to the combination of audio and spatially and temporally coherent sensory-motor information. Nevertheless, even mere exposure to sound can change acoustic localization[12,13] because other proprioceptive information is typically used to localize sounds, such as head position and movements.

Here, we analyzed the plasticity of human spatial hearing by considering different sources of errors in sound localization, namely the *RMSE*, *gain*, *bias*, and *dispersion* (i.e., the inverse of the precision). Our manipulation of the spectral cues dramatically affected sound localization for all such indexes except *bias*, and so did our training protocol: while wearing the new ears, performance dropped across the board, and similarly, all sources of errors dropped to near-baseline levels in participants that underwent training, but not in the control group. This demonstrates that audio-proprioceptive training jointly reduces systematic and random localization errors.

Although in most domains of perception, adaptation is followed by aftereffects,[44] previous studies failed to provide evidence for aftereffects in spatial hearing following adaptation to novel spectral cues.[12,31] By adapting participants to large HRTF distortions and using analyses capable of isolating different sources of localization errors, this study provides concluding evidence for the existence of systematic aftereffects in spatial hearing induced by the adaptation to novel HRTFs. Indeed, aftereffects only occurred for participants who underwent the training and successfully learned to localize sounds with the new ears. The analysis performed on the HRTFs measured with the original and new ears highlighted that the manipulation performed did not cause a simple alteration of the spectral cues (as done in previous studies) but rather a global disruption of the spectral cues. This may explain why previous studies failed to find analogous aftereffects. Notably, such aftereffects disappeared within a few trials upon removal of the new ears (Figure 4) and were only evident in the random error component (an aspect that could have passed unnoticed in previous studies, which relied on less sensitive analyses of localization errors). This demonstrates that while listeners can quickly switch between different HRTFs, such a switch nevertheless comes at a cost.

Nevertheless, to our knowledge, only one work by Trapeau et al.[14] addressed the question of the absence of aftereffect following adaptation to changing spectral cues (application of molds), even in localizing sounds vertically. Despite the training lasting six days (as opposed to ours lasting 20 min), the authors found that participants returned to localizing sounds accurately after removal of the molds in all groups tested (independently from the kind of training) with no aftereffect observed even looking at the effect in a trial-by-trial analysis. Despite the different types and durations of the training, these results are in line with our results except for the aftereffects of the dispersion. Indeed, while we found an increase in the aftereffect index in the training group (and not in the controls), Trapeau et al.[14] found this effect only in the control group. However, the presence of this effect by the authors was not discussed, which proves that this aspect may have gone unnoticed in previous studies.

The present study highlights the plastic nature of human spatial hearing and demonstrates that our auditory system can rapidly adapt to changes in spectral cues, given appropriate training. Learning to use spectral cues for spatial hearing efficiently is especially relevant for virtual auditory spaces (which often rely on generic HRTFs), hearing aids, and early blind individuals, whose vertical sound localization is usually impaired[45,46]; this study provides a handle to quickly train and assess human spatial hearing along the vertical plane.

Limitations of the study

The present study quantified how the auditory system can rapidly adapt to spectral cue change following sensory motor training. A limitation of this study is that we did not further investigate the correlation between the behavioral effects of training and HRTF of individual participants. This point also relates to the small sample undergoing sensorimotor training and the control groups. The small sample size would have reduced the statistical power and reliability of the results. Therefore, future studies are needed to investigate this point.

Furthermore, as a suggestion for future studies, all the speakers could be equidistant from the participants, arranging them in a semicircle and not vertically as in our study. Keeping the speakers equidistant would make the paradigm "neat", so the variation in volume and pitch between the outer and middle speakers could not be considered confounding in absolute localization performance. However, we would like to point out that the effect of the new ears should act on spectral cues and not intensity cues.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Monica Gori (monica.gori@iit.it).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The data have been deposited on Zenodo (link: https://zenodo.org/records/8374758) and are publicly available as of the date of publication.
- 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.

ACKNOWLEDGMENTS

This work was supported by the project MySpace (g.a. 948349), founded by the European Research Council and by the European Union Horizon Europe Research and Innovation program under the Marie Skłodowska-Curie-2021-PF-01 FLEX-U (g.a. 101064748).

AUTHOR CONTRIBUTIONS

Conceptualization: C.P., M.G., S.F., and M.E.; methodology: C.P., M.G., S.F., and M.E.; data curation: S.F., D.E., and A.T.; formal analysis: A.T.; software: A.T.; validation: A.T.; visualization: A.T. and C.P.; writing—original draft preparation: A.T. and C.P.; writing—review and editing: M.E., M.G., D.E., and S.F.; funding acquisition: M.G.

DECLARATION OF INTERESTS

The authors declare no conflicting interests.

STAR★METHODS

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

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.111308.

Received: November 27, 2023
Revised: April 15, 2024
Accepted: October 30, 2024
Published: November 4, 2024

REFERENCES

1. Parks, T.N., Rubel, E.W., Popper, A.N., and Fay, R.R. (2004). Plasticity of the Auditory System (Springer). https://doi.org/10.1007/978-1-4757-4219-0.
2. Gilbert, C.D., Sigman, M., and Crist, R.E. (2001). The Neural Basis of Perceptual Learning. Neuron *31*, 681–697. https://doi.org/10.1016/S0896-6273(01)00424-X.
3. King, A.J., Dahmen, J.C., Keating, P., Leach, N.D., Nodal, F.R., and Bajo, V.M. (2011). Neural circuits underlying adaptation and learning in the perception of auditory space. Neurosci. Biobehav. Rev. *35*, 2129–2139. https://doi.org/10.1016/j.neubiorev.2011.03.008.
4. Mendonça, C. (2014). A review on auditory space adaptations to altered head-related cues. Front. Neurosci. *8*, 219. https://doi.org/10.3389/fnins.2014.00219.
5. Wright, B.A., and Zhang, Y. (2009). A review of the generalization of auditory learning. Philos. Trans. R. Soc. Lond. B Biol. Sci. *364*, 301–311. https://doi.org/10.1098/rstb.2008.0262.
6. Javer, A.R., and Schwarz, D.W. (1995). Plasticity in human directional hearing. J. Otolaryngol. *24*, 111–117.
7. Slattery, W.H., and Middlebrooks, J.C. (1994). Monaural sound localization: Acute versus chronic unilateral impairment. Hear. Res. *75*, 38–46. https://doi.org/10.1016/0378-5955(94)90053-1.
8. Carlile, S., and Blackman, T. (2014). Relearning Auditory Spectral Cues for Locations Inside and Outside the Visual Field. J. Assoc. Res. Otolaryngol. *15*, 249–263. https://doi.org/10.1007/s10162-013-0429-5.
9. Majdak, P., Walder, T., and Laback, B. (2013). Effect of long-term training on sound localization performance with spectrally warped and band-limited head-related transfer functions. J. Acoust. Soc. Am. *134*, 2148–2159. https://doi.org/10.1121/1.4816543.
10. Mendonça, C., Campos, G., Dias, P., and Santos, J.A. (2013). Learning Auditory Space: Generalization and Long-Term Effects. PLoS One *8*, e77900. https://doi.org/10.1371/journal.pone.0077900.
11. Parseihian, G., and Katz, B.F.G. (2012). Rapid head-related transfer function adaptation using a virtual auditory environment. J. Acoust. Soc. Am. *131*, 2948–2957. https://doi.org/10.1121/1.3687448.
12. Hofman, P.M., Van Riswick, J.G., and Van Opstal, A.J. (1998). Relearning sound localization with new ears. Nat. Neurosci. *1*, 417–421. https://doi.org/10.1038/1633.
13. Florentine, M. (1976). Relation between lateralization and loudness in asymmetrical hearing losses. J. Am. Audiol. Soc. *1*, 243–251.
14. Trapeau, R., Aubrais, V., and Schönwiesner, M. (2016). Fast and persistent adaptation to new spectral cues for sound localization suggests a many-to-one mapping mechanism. J. Acoust. Soc. Am. *140*, 879–890. https://doi.org/10.1121/1.4960568.
15. Irving, S., and Moore, D.R. (2011). Training sound localization in normal hearing listeners with and without a unilateral ear plug. Hear. Res. *280*, 100–108. https://doi.org/10.1016/j.heares.2011.04.020.
16. Shams, L., and Seitz, A.R. (2008). Benefits of multisensory learning. Trends Cognit. Sci. *12*, 411–417. https://doi.org/10.1016/j.tics.2008.07.006.
17. Berger, C.C., Gonzalez-Franco, M., Tajadura-Jiménez, A., Florencio, D., and Zhang, Z. (2018). Generic HRTFs May be Good Enough in Virtual Reality. Improving Source Localization through Cross-Modal Plasticity. Front. Neurosci. *12*, 21. https://doi.org/10.3389/fnins.2018.00021.
18. Valzolgher, C., Campus, C., Rabini, G., Gori, M., and Pavani, F. (2020). Updating spatial hearing abilities through multisensory and motor cues. Cognition *204*, 104409. https://doi.org/10.1016/j.cognition.2020.104409.
19. Strelnikov, K., Rosito, M., and Barone, P. (2011). Effect of Audiovisual Training on Monaural Spatial Hearing in Horizontal Plane. PLoS One *6*, e18344. https://doi.org/10.1371/journal.pone.0018344.
20. Shinn-Cunningham, B.G., Durlach, N.I., and Held, R.M. (1998). Adapting to supernormal auditory localization cues. II. Constraints on adaptation of

mean response. J. Acoust. Soc. Am. *103*, 3667–3676. https://doi.org/10.1121/1.423107.

21. Cuturi, L.F., and Macneilage, P.R. (2014). Optic flow induces nonvisual self-motion aftereffects. Curr. Biol. *24*, 2817–2821. https://doi.org/10.1016/j.cub.2014.10.015.
22. Domenici, N., Tonelli, A., and Gori, M. (2022). The development of adaptation aftereffects in the vibrotactile domain. J. Exp. Psychol. Gen. *151*, 3134–3143. https://doi.org/10.1037/xge0001252.
23. Tonelli, A., Pooresmaeili, A., and Arrighi, R. (2020). The Role of Temporal and Spatial Attention in Size Adaptation. Front. Neurosci. *14*, 539. https://doi.org/10.3389/fnins.2020.00539.
24. Senna, I., Parise, C.V., and Ernst, M.O. (2017). Modulation frequency as a cue for auditory speed perception. Proc. Biol. Sci. *284*, 20170673. https://doi.org/10.1098/rspb.2017.0673.
25. Vroomen, J., and de Gelder, B. (2003). Visual Motion Influences the Contingent Auditory Motion Aftereffect. Psychol. Sci. *14*, 357–361. https://doi.org/10.1111/1467-9280.24431.
26. Schweinberger, S.R., Casper, C., Hauthal, N., Kaufmann, J.M., Kawahara, H., Kloth, N., Robertson, D.M.C., Simpson, A.P., and Zäske, R. (2008). Auditory Adaptation in Voice Perception. Curr. Biol. *18*, 684–688. https://doi.org/10.1016/j.cub.2008.04.015.
27. Ben Porquis, L., Finocchietti, S., Zini, G., Cappagli, G., Gori, M., and Baud-Bovy, G. (2017). ABBI: A wearable device for improving spatial cognition in visually-impaired children. In 2017 IEEE Biomedical Circuits and Systems Conference (BioCAS) (IEEE), pp. 1–4. https://doi.org/10.1109/BIOCAS.2017.8325128.
28. Finocchietti, S., Cappagli, G., and Gori, M. (2015). Encoding audio motion: spatial impairment in early blind individuals. Front. Psychol. *6*, 1357. https://doi.org/10.3389/fpsyg.2015.01357.
29. Cappagli, G., Finocchietti, S., Cocchi, E., Giammari, G., Zumiani, R., Cuppone, A.V., Baud-Bovy, G., and Gori, M. (2019). Audio motor training improves mobility and spatial cognition in visually impaired children. Sci. Rep. *9*, 3303. https://doi.org/10.1038/s41598-019-39981-x.
30. Cuppone, A.V., Cappagli, G., and Gori, M. (2018). Audio Feedback Associated With Body Movement Enhances Audio and Somatosensory Spatial Representation. Front. Integr. Neurosci. *12*, 37. https://doi.org/10.3389/fnint.2018.00037.
31. Steadman, M.A., Kim, C., Lestang, J.-H., Goodman, D.F.M., and Picinali, L. (2019). Short-term effects of sound localization training in virtual reality. Sci. Rep. *9*, 18284. https://doi.org/10.1038/s41598-019-54811-w.
32. Tonelli, A., Brayda, L., and Gori, M. (2015). Task-dependent calibration of auditory spatial perception through environmental visual observation. Front. Syst. Neurosci. *9*, 84. https://doi.org/10.3389/fnsys.2015.00084.
33. Lovelace, E.A., and Anderson, D.M. (1993). The role of vision in sound localization. Percept. Mot. Skills *77*, 843–850.
34. Alais, D., and Burr, D.C. (2004). Ventriloquist Effect Results from Near-Optimal Bimodal Integration. Curr. Biol. *14*, 257–262. https://doi.org/10.1016/S0960-9822(04)00043-0.
35. Bizley, J.K., and King, A.J. (2008). Visual–auditory spatial processing in auditory cortical neurons. Brain Res. *1242*, 24–36.
36. Atilgan, H., Town, S.M., Wood, K.C., Jones, G.P., Maddox, R.K., Lee, A.K.C., and Bizley, J.K. (2018). Integration of Visual Information in Auditory Cortex Promotes Auditory Scene Analysis through Multisensory Binding. Neuron *97*, 640–655.e4. https://doi.org/10.1016/j.neuron.2017.12.034.
37. Froese, T., and González-Grandón, X. (2020). How passive is passive listening? Toward a sensorimotor theory of auditory perception. Phenomenol. Cognitive Sci. *19*, 619–651. https://doi.org/10.1007/s11097-019-09641-6.
38. Lohse, M., Zimmer-Harwood, P., Dahmen, J.C., and King, A.J. (2022). Integration of somatosensory and motor-related information in the auditory system. Front. Neurosci. *16*, 1010211. https://doi.org/10.3389/fnins.2022.1010211.
39. Caclin, A., Soto-Faraco, S., Kingstone, A., and Spence, C. (2002). Tactile "capture" of audition. Percept. Psychophys. *64*, 616–630. https://doi.org/10.3758/BF03194730.
40. Murray, M.M., Molholm, S., Michel, C.M., Heslenfeld, D.J., Ritter, W., Javitt, D.C., Schroeder, C.E., and Foxe, J.J. (2005). Grabbing Your Ear: Rapid Auditory–Somatosensory Multisensory Interactions in Low-level Sensory Cortices Are Not Constrained by Stimulus Alignment. Cerebr. Cortex *15*, 963–974. https://doi.org/10.1093/cercor/bhh197.
41. Reznik, D., and Mukamel, R. (2019). Motor output, neural states and auditory perception. Neurosci. Biobehav. Rev. *96*, 116–126. https://doi.org/10.1016/j.neubiorev.2018.10.021.
42. Aytekin, M., Moss, C.F., and Simon, J.Z. (2008). A sensorimotor approach to sound localization. Neural Comput. *20*, 603–635. https://doi.org/10.1162/neco.2007.12-05-094.
43. Bernard, M., Pirim, P., de Cheveigne, A., and Gas, B. (2012). Sensorimotor learning of sound localization from an auditory evoked behavior. In 2012 IEEE International Conference on Robotics and Automation (IEEE), pp. 91–96. https://doi.org/10.1109/ICRA.2012.6224972.
44. Wark, B., Lundstrom, B.N., and Fairhall, A. (2007). Sensory adaptation. Curr. Opin. Neurobiol. *17*, 423–429. https://doi.org/10.1016/j.conb.2007.07.001.
45. Lewald, J. (2002). Vertical sound localization in blind humans. Neuropsychologia *40*, 1868–1872. https://doi.org/10.1016/S0028-3932(02)00071-4.
46. Voss, P., Tabry, V., and Zatorre, R.J. (2015). Trade-Off in the Sound Localization Abilities of Early Blind Individuals between the Horizontal and Vertical Planes. J. Neurosci. *35*, 6051–6056. https://doi.org/10.1523/JNEUROSCI.4544-14.2015.
47. Tabry, V., Zatorre, R.J., and Voss, P. (2013). The influence of vision on sound localization abilities in both the horizontal and vertical planes. Front. Psychol. *4*, 932. https://doi.org/10.3389/fpsyg.2013.00932.
48. The MathWorks Inc (2022). MATLAB version: 9.13.0 (R2022b). https://au.mathworks.com/help/matlab/ref/rmse.html.
49. The MathWorks Inc (2022). Curve Fitting Toolbox Version 3.8 (R2022b). https://au.mathworks.com/help/curvefit/fit.html.
50. Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q.F., Smíra, M., Epskamp, S., et al. (2019). JASP : Graphical Statistical Software for Common Statistical Designs. J. Stat. Software *88*, 1–17. https://doi.org/10.18637/jss.v088.i02.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Dataset generated for the experiment	Zenodo	https://zenodo.org/records/8374758
Software and algorithms		
G*Power version 3.1.9.4	hhu	https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower.html
MATLAB R2022b	MathWorks	https://www.mathworks.com/
JASP version 0.19.0.0	JASP	https://jasp-stats.org/

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Experiment 1 participants

We recruited thirty-four participants (twenty females, with an average age of 28.76, *SD* = 4.57). All participants had normal or corrected to normal vision and normal hearing. They gave written informed consent before starting the experiment. The local health service ethics committee approved the study (Comitato Etico Regione Liguria, Genoa, Italy; Prot. IIT_UVIP_COMP_2019 N. 02/2020, 4 July 2020) and followed the Declaration of Helsinki.

We conducted an *a priori* power analysis using G*Power version 3.1.9.4 to decide the sample size. Considering a medium effect of 0.5 with 0.8 of power, at a significance criterion of $\alpha = 0.05$, we obtain a sample size of 34 (difference between two dependent means, matched pairs). Thus, the obtained sample size of $N = 34$ is adequate to test the study hypothesis.

Experiment 2 participants

For Experiment 2, we randomly selected 16 of the participants who also participated in Experiment 1. (nine females, with an average age of 31.06, SD = 4.42).

METHOD DETAILS

Experiment 1 stimuli and setup

The sound stimuli were generated with a custom-built array of speakers (Figure 1A), controlled using MATLAB. The array of speakers was composed of 10 loudspeakers (total height: 75 cm) placed vertically on a small table so that the center was positioned at the participant's head height. The loudspeakers were placed at a distance of 40 cm from the listener; such a short distance was chosen to match the range of distances used during the training and facilitate pointing responses. Acoustic stimuli consisted of bursts of white noise with a rectangular envelope lasting 1 s. Note that due to the arrangement of the loudspeakers on a straight line, different loudspeakers might have colored the white noise differently due to their directivity pattern.

Experiment 1 procedure

Before entering the experiment room, participants were blindfolded to prevent seeing the loudspeakers from affecting their performance[32,47] and remained blindfolded throughout. Apart from this practical reason, blindfolding participants enables (1) testing auditory plasticity in a more challenging condition, as vision provides rich spatial cues for sound localization and (2) easily extending the procedure to visually-impaired populations. At the beginning of the experiment, participants laid their heads on a chinrest to ensure that the head position remained fixed and aligned with the speaker's array's center for the duration of the experiment. Participants sat in front of the speakers, and their ' task was to touch the speakers from which the sound came while the experimenter read the position 'from a measuring tape attached to the side of the speaker array.

After a short break of ∼10 min, each participant repeated the task, wearing the new ears to modify the head-related transfer function (Figure 1B). Each position was repeated five times for a total of 50 trials for each session.

Experiment 2 stimuli and setup

We used the same stimuli as Experiment 1 (Figure 1A) for the localization task. Additionally, a wrist-worn wearable, custom-designed device with an integrated audio system and inertial sensors,[27] was used for audio-proprioceptive training. This device couples the movements of the users' arms with spatiotemporally co-localized acoustic feedback. For the experiment, the device was set to produce white noise (70 dB SPL) acoustic feedback, and the sound was activated as soon as the arm was moved, thanks to inertial sensors in the device (for technical information, see[27]).

Experiment 2 procedure

Participants were divided into two groups: the training and the control group. As illustrated in Figure 2A, both groups repeated five times the same vertical auditory localization task used in experiment 1 (see Experiment 1's methods subsection for more details). In the first and last sessions (*regular ears 1* and *regular ears 2*), all participants performed the task without the new ears, while in the central sessions (*new ears 1, new ears 2, new ears 3*) they wore them.

Unlike the control group, the training group used the bracelet device for a short training before the *new ears 2* and *new ears 3* sessions. During the training, participants wore the bracelet on their wrist and moved it along the array of speakers without touching it. At the same time, the device produced a white noise synchronized with the participant's movement for ∼10 min. There was no constraint on the arm movement, including the movement of the elbow.

Conversely, the control group made the same movements along the speaker array but without any feedback sound.

As in experiment 1, each position was repeated five times for a total of 50 trials for each session. The experiment lasted ∼90 min.

QUANTIFICATION AND STATISTICAL ANALYSIS

Experiment 1 data analysis

To analyze the data for each participant, we calculated four dependent variables: the root-mean-squared-error (*RMSE*), the *gain*, the *bias,* and the *dispersion*. The first variable, the *RMSE*, provides an overall measure of localization performance, including random and constant error terms. We computed the *RMSE* using the MATLAB function called *rmse*.[48]

Next, using the MATLAB "Curve fitting toolbox",[49] we calculated the linear fit between the perceived position and the real position of the sound using the function *fit (x, y,' poly1')* for each participant. From the fit, we obtained *gain*, *bias,* and *dispersion*. The *gain* was calculated as the slope of the regression line; a *gain* below one indicates compression of perceived elevation, whereas a *gain* above one indicates an expansion of perceived elevation. The *bias* was calculated as the intercept of the fit, which gives an estimate of constant errors at the central speaker. The *dispersion* describes the random error and corresponds to sensorimotor precision (the lower the dispersion, the higher the precision).

All t-tests were done using JASP software[50] and Cohen's *d* gives the effect size.

Experiment 2 data analysis

As for experiment 1, we calculated the *RMSE* and the best linear fit between the perceived position and the real position of the sound, obtaining the gain, the bias, and the dispersion. This procedure was carried out for each session and participant.

Using the *RMSE*, *gain*, *bias,* and *dispersion* as dependent variables, we ran three separate 5x2 repeated measure ANOVAs with between factor *Group* (training, control) and within factor session *(regular ears 1, new ears 1, new ears 2, new ears 3, regular ears 2).* As a follow-up post hoc analysis, we used paired or unpaired two-tail Student's t-tests, depending on whether the comparison was among variables between or within groups. For the complete report of results, see the Supplementary materials.

Furthermore, we calculated two new indices for each dependent variable to better quantify the effect of both the manipulation of the HRTFs and the acoustic-proprioceptive training: the "*Learning Index*" and the "*Aftereffect Index*" (Figure 3).

The "*Learning Index*" is calculated for each dependent variable by subtracting its value in the session *new ears 1* from that of the session *new ears 3*. We repeated this operation for each participant. This index quantifies the improvement in localization performance following training.

To assess whether adaptation to novel spatial cues induced an aftereffect, we calculated the "*Aftereffect Index*" for each dependent variable by subtracting its value in the session *regular ears 1* from that of the session *regular ears 2*. We repeated this operation for each participant.

To check whether the two indexes differed between the groups, we used an unpaired two-tailed t-test.

As a final analysis, to check the duration of the aftereffect, we divided the trials of the last session (*Regular Ears 2*) by taking the first 25 trials (*First Half*) and the last 25 trials (*Second Half*), and for each participant, we calculated the linear fit between the perceived position and the real position for each of the two blocks. We ran 2x2 repeated measure ANOVAs with between factor *Group* (training, control) and within factor *Block* (*First Half* and *Second Half*), and as post hoc analysis, we used paired or unpaired two-tail Student's t-tests.

All *p*-values for the post hoc t-test were corrected for multiple comparisons (Bonferroni), and Cohen's *d* gives the effect size. All ANOVAs and t-tests were done using JASP software.[50].