

# The link between blindness onset and audiospatial processing: testing audiomotor cues in acoustic virtual reality

Davide Esposito, Alice Bollini, and Monica Gori\*

**Abstract**— Vision seems essential for cross-modal calibration of auditory cues in spatial perception. Previous findings showed that, in some specific tasks such as sound localization, blind individuals have enhanced skills, suggesting that the audiomotor loop might partially compensate for early visual loss in the calibration of auditory space; however, direct evidence is still lacking. Here, we proposed a method based on the alteration of the audiomotor loop. Acoustic virtual reality was used to measure the audiomotor loop's influence on the space perception of blind individuals. We developed a VR steering task by head or trunk pointing to auditory sources, where the audiomotor conflict is induced by letting trunk rotations change the auditory scene together with head rotations. Early blind, late blind, and sighted participants were tested to assess their sensitivity to the induced audiomotor conflict. The platform demonstrated its effectiveness in exposing participants' sensitivity to the audiomotor loop alteration. The early blind group was significantly more affected than the sighted group, while the late blind group did not significantly differ from any of the other groups. Our results confirm the increased role of the audiomotor loop for audiospatial information processing in blindness and advocate for the development of new spatial orientation training for blind people based on exploiting the audiomotor loop itself.

## I. INTRODUCTION

Scientific results suggest that spatial cognition is acquired during development [1]. In the context of auditory space perception, the importance of vision for the calibration of audio-spatial cues is well known [2]. Indeed, several aspects of spatial performance are impaired when vision is lost early in life [3]. However, early and late blind showed from sighted-like to enhanced skills in azimuthal sound localization and orientation [3], suggesting the existence of a compensatory process for audio-spatial calibration in absence of vision. Albeit not proved yet, it was proposed that the audiomotor loop, that is, the association between movement and consequent change of auditory scene, may drive the compensatory calibration in early blind people [4]. If that were the case, early blind people would use the audiomotor loop to encode audio-spatial information. Understanding the sensorimotor processes visually deprived people undergo to calibrate their auditory space is paramount for developing mobility and orientation training suited for their condition [5].

Notwithstanding that, to the best of our knowledge, a standard method to assess the audiomotor loop's contribution to audio-spatial processing does not exist. Here, we present a method based on the inducement of an audiomotor conflict. Our approach is based on the hypothesis that people whose auditory space was calibrated by the audiomotor loop would be more affected by an alteration of the loop itself. To induce the audiomotor conflict, we have developed an ad-hoc

platform based on auditory virtual reality in the shape of an archery-like game. Our auditory virtual reality platform altered the audiomotor loop by rotating the virtual ears according to the head and trunk rotations. Early blind, late blind, and sighted people were tested on the VR platform. If the platform could effectively alter the audiomotor loop, and early blind people were the only ones who relied on the audiomotor loop to calibrate their auditory space, they would experience a more severe performance drop than late blind and sighted people.

## II. METHODS

### A. Participants

In total, 27 individuals, 10 early blind (3 males, 7 females, age =  $33.2 \pm 3.19$  years old), 6 late blind (3 males, 3 females, age =  $45.67 \pm 5.56$ ) and 11 sighted individuals (6 males, 5 females, age =  $31.27 \pm 3.92$  years old) were involved in the study. Based on the literature on audiovisual calibration [6], blind people were considered as early if they lost vision before the age of 5, whereas late blind people lost vision after the age of 10. The clinical details of the participants' pathologies are reported in table 1. The early and late blind group sizes amounted to the number of contacts that we managed to recruit in one year. All of them were enrolled from the local contacts of Genoa. Informed consent was obtained from all participants. The study followed the tenets of the Declaration

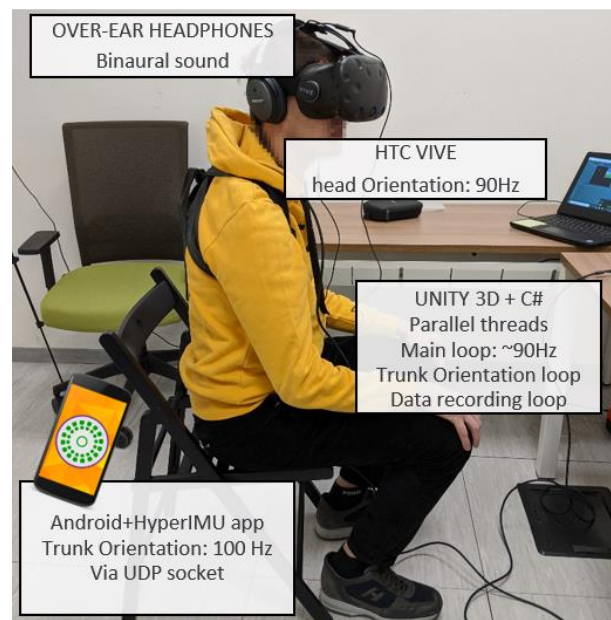


Fig. 1: experimental setting. The user sits on a chair without resting his/her back, which is free to move. The trunk sensor is kept in position by a custom-made harness. The user wears the head-mounted display and a pair of headphones.

\* D.E. is with Istituto Italiano di Tecnologia, Via E. Melen 83, GE 16152 Italy, and with DIBDRIS, Università di Genova, Via All'Opera Pia, 13, GE 16145 Italy (corresponding author to provide phone: +39 0108172232; e-mail: davide.esposito@iit.it).

A.B. is with Istituto Italiano di Tecnologia (e-mail: alice.bollini@iit.it).  
M.G. is with Istituto Italiano di Tecnologia (e-mail: monica.gori@iit.it).

TABLE I. BLIND PARTICIPANTS' CLINICAL DETAILS

	Gender	Age	Pathology	Blindness onset	Residual vision
E1	F	32	Retinopathy	Before birth	No vision
E2	F	20	Retinopathy	Before birth	Lights and shadows
E3	F	29	Retinopathy	Before birth	No vision
E4	M	27	Leber's amaurosis	Since birth	No vision
E5	F	26	Glaucoma and retinal detachment	Before birth	No vision
E6	M	46	Leber's disease	Before birth	No vision
E7	M	52	Unknown	Before birth	Lights and shadows
E8	F	30	Retinitis pigmentosa	Since birth	Lights and shadows
E9	F	28	Microphthalmia	Since birth	No vision
E10	F	42	Retinopathy	Since birth	No vision
L1	M	62	Uveitis	11 years	Lights and shadows
L2	M	29	Leber's Amaurosis	13 years	Lights and shadows
L3	M	32	Corneal opacity	17 years	No vision
L4	F	41	Retinitis pigmentosa & Stalगत Syndrome	30 years	Lights and shadows
L5	F	55	Retinitis pigmentosa & Nystagmus	30 years	Lights and shadows
L6	F	55	Leber's amaurosis	46 years	No vision

of Helsinki and was approved by the ethics committee of the local health service (Comitato Etico, ASL 3, Genova).

### B. Experimental setting

The experiments were conducted in a quiet room. Participants were sitting on a chair; they could rest their back only during breaks, while during the experiment, they were asked to sit close to the chair edge and their trunk was free to rotate. The sitting position was chosen for safety reasons since we did not know whether the audiomotor loop alteration could impair participants' equilibrium. The virtual environment created for this experiment was developed with the game engine Unity 3D. The spatial blending of sounds was made using the *resonance* package [7]. The sound was delivered via commercially available BOSE® over-ear headphones. The head-mounted display (HMD) visual output was blanked, so the virtual reality headset (HTC® VIVE) was used uniquely to track the participants' head movements. The participants' head rotations were tracked by the HMD itself, which provides 6-DOF orientation and position data at a sampling rate of 90 Hz, which is the frequency of Unity's main loop. The trunk rotations were tracked by an LG® google nexus 4 smartphone used as a wireless inertial measurement unit thanks to the app HyperIMU [8], which provides 3-DOF orientation data at a nominal sampling rate of 100 Hz. Both sensors have a resolution of  $0.1^\circ$ . The incoming samples were asynchronously collected via a UDP socket, stacked and averaged at the frequency of 90 Hz. The smartphone was fixed to their back using a custom-made harness (Fig. 1).

The virtual environment is designed as an archery field. Unity's unit of measure for length is the unity unit; however, objects' proportions matched the unity unit with the meter. Hence the meter will be used to describe the spatial parameters. The camera view is 1.7 m above the ground. The target is 60 m distant from the starting point and emits an intermittent pink noise with a 5 Hz duty cycle and isotropic logarithmic spatial attenuation. The target can appear in three positions concerning the participant's seat:  $-15^\circ$ ,  $0^\circ$ ,  $+15^\circ$ . The target positions were chosen as such because they provided an appropriate degree of difficulty in pilot tests during the platform development. Yet, those values can be easily modified according to the experimenter's requirements.

### C. Task description

The auditory virtual reality platform implements a goal-directed steering task in form of an archery-like game, where participants are told to be on the arrow and are requested to drive themselves towards the 3D rendered virtual sound source. The arrow shooting position is always turned approximately  $15^\circ$  rightwards or leftwards from the target position. See Fig. 2A for a description of the four possible geometrical relationships between arrow and target. The arrow is automatically shot as the participant keeps it for 1 to 3 s in the given shooting position range, that is  $\theta_{\text{target}} \pm 15^\circ + [-3, 3]^\circ$ , randomized across trials. Before the arrow is shot, a pitch-modulated intermittent pure tone feeds distance from the shooting position back to the participants to help them find the starting orientation. Once shot, the arrow moves at 10 m/s constant velocity; its movement is constrained to the horizontal plane only, and participants can control its travel direction only. The trial ends when the arrow hits the target or the virtual wall aligned with the target. The sound of an arrow hitting the target is played back when the target is hit; contrarily, the sound of a swish is played back.

Two first-person driving modalities have been designed according to the state of the audiomotor loop, defined by the contribution of head and trunk rotations to the rotation of arrow and virtual ears (i.e., rotation of the virtual listener, which rules the perceived angular distance from the audio source). In the baseline condition, when the audiomotor loop is *preserved*, both arrow and virtual ears mimic the head yaw. Hence the task is simple steering in first-person perspective. In this condition, if during one trial participants had, say, head and trunk both rotated  $15^\circ$  to the right, the arrow direction would be  $15^\circ$  to the right (head yaw), and the sound source would be heard as coming from  $15^\circ$  to the left of the ears (head yaw). In the test condition, when the audiomotor loop is *altered*, the arrow mimics the trunk yaw and the virtual ears mimic the sum of head and trunk yaws. Taking as example the same scenario as before, the arrow direction would be  $15^\circ$  to the right (trunk yaw), while the sound source would be heard as coming from  $30^\circ$  to the left of the ears (sum of head and trunk yaw). Fig. 2B graphically describes the relationship between head, trunk, virtual ears, and arrow for each driving

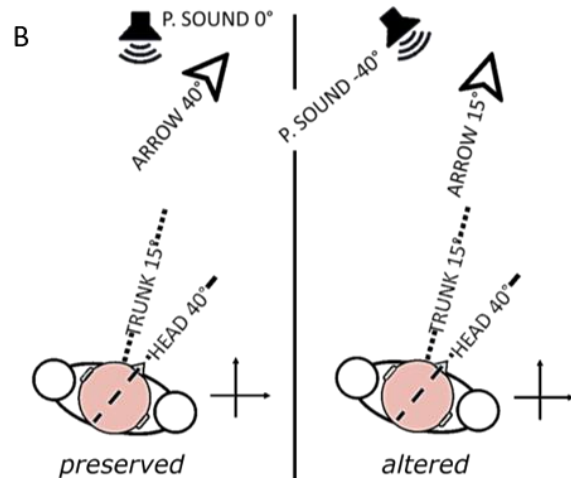
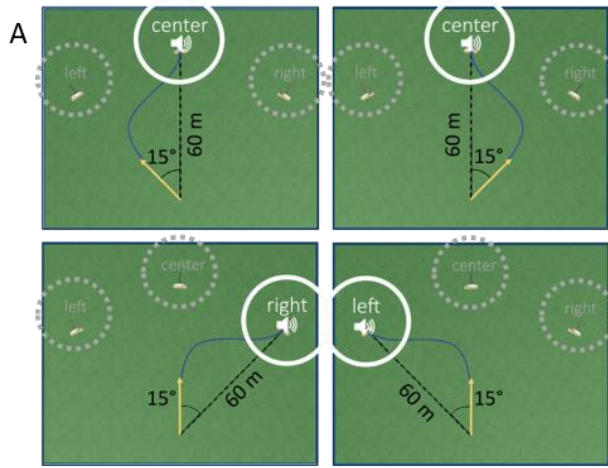


Fig. 2: In panel A, grid of the four possible geometrical relationships between target and initial arrow orientations in the virtual environment. Targets circled in white are active, those circled in grey are inactive. The yellow arrows describe the initial arrow orientations. The blue curved lines show plausible arrow trajectories in one trial. In panel B, example describing the relationship between head, trunk, virtual ears and arrow for the two “coordination” levels.

modality. In addition to that, in trials with *altered* audiomotor loop, participants were required to immobilize their head to the absolute straight-ahead as much as they could.

#### D. Experimental procedure

Firstly, participants were given the following instructions: "Imagine you are on an arrow. Once shot, it will fly at a constant speed and you will control its direction only by moving a part of your body, as I will tell you. Your goal is to drive the arrow towards the sound you will hear, which will correspond to the target. There are three possible targets." Then, they were blindfolded and introduced to the virtual platform by the tactile exploration of a scaled plastic model of the environment, and the experimenter made them track with the finger four plausible arrow paths, one for each possible initial arrow-target geometrical relationship (i.e., the path drawn by the curved blue lines in Fig. 2A). To let participants familiarize themselves with the platform, a very short training session was performed, made of no more than eight trials in the *preserved* condition, and eight trials in the *altered* condition.

The trials with the target at  $0^\circ$  were separated from those with the target at  $\pm 15^\circ$  to reduce the eventuality of ending the trial in the next trial's shooting range. Therefore, the experiment consisted of four blocks, corresponding to the VR platform's two driving modalities and the two couples of arrow-target configurations. They were performed consecutively, in randomized order, and counterbalanced among participants to prevent learning from affecting the consecutive statistical analysis. Each block was made of twenty trials. Immediately before starting a run, the participant was explained by the experimenter the requirements for that run's condition. The experiment lasted approximately 30 minutes with no breaks. Breaks were allowed anytime according to participants' needs. Fig. 3 schematically describes the experimental design as a flowchart.

#### E. Data analysis

The auditory virtual reality platform provides kinematic data for the head, trunk, and virtual arrow. Instantaneous head orientation, trunk orientation, and arrow position are stored for each trial and available for further analysis. For the sake of the

current scientific question, only the distributions of final arrow positions relative to the targets were analyzed, split by driving modality. Data for each subject and each driving modality were aggregated by means of the mean absolute error (MAE), the error being the linear distance between target center and arrow endpoint. The MAE indicates participants' ability to control the arrow. MAE distributions' normality was assessed via Shapiro-Wilk test, which was rejected for every MAE distribution. Therefore, subsequent analyses were performed using non-parametric methods. Specifically, MAE samples underwent a two-ways  $3 \times 2$  mixed ANOVA on aligned and rank transformed data, or ART ANOVA [9], with "group" as a between-subjects factor and "audiomotor loop state" as a within-subjects factor. Post-hoc differences among groups were assessed via Mann-Whitney test; within-group

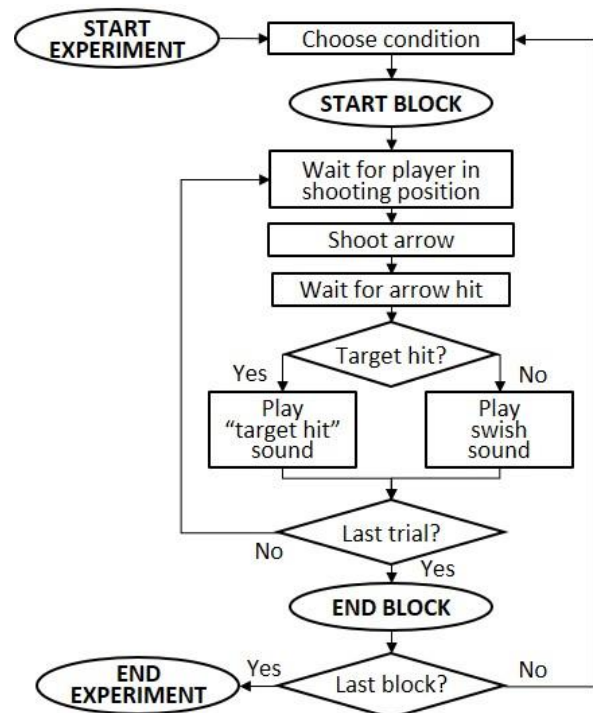


Fig. 3: Flowchart of the experimental design.

differences were assessed via Wilcoxon test. The comparisons were Bonferroni corrected. All the analyses were performed with the software R. Datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

### III. RESULTS

If early blind people were more affected by the audiomotor loop alteration, they would show larger MAE values than late blind and sighted people with *altered* audiomotor loop, while they would not show any significant difference with *preserved* audiomotor loop. The ART ANOVA resulted significant for “audiomotor loop state”,  $F(1,25)=80.38$ ,  $\eta_p^2=0.76$ ,  $p<0.001$ , but not for “group”  $F(2,25)=2.98$ ,  $\eta_p^2=0.19$ ,  $p=0.069$ . Their interaction resulted significant,  $F(2,25)=6.98$ ,  $\eta_p^2=0.36$ ,  $p=0.004$ . The post-hoc Mann-Whitney tests showed no difference held between groups when the audiomotor loop was *preserved*,  $p\geq 0.770$  for all the comparisons. Instead, when the audiomotor loop was *altered*, the early blind group's MAEs were significantly larger than those of the sighted group,  $p=0.023$ , but not significantly different from those of the late blind group,  $p=0.108$ , which in turn did not differ from those of the sighted group,  $p=1.000$ . The post-hoc Wilcoxon tests showed the MAE was significantly larger when the audiomotor loop was *altered* than when it was *preserved*,  $p_{early}=0.027$ ,  $p_{late}=0.027$ ,  $p_{sighted}=0.019$ , See Fig. 4 for a graphic representation of MAE distributions and post-hoc results.

### IV. DISCUSSION

This study aimed at assessing the relationship between blindness onset and use of the audiomotor loop for the calibration of the horizontal audio space [1], [3]. To do so, 10 early blind, 6 late blind and 11 sighted people were tested on a goal-directed steering task in auditory virtual reality where their audiomotor loop could be preserved or altered. Early blind people lacked vision during the critical ages for cross-modal calibration ( , 5 to 10 [6], [10]), whereas late blind and sighted people had vision during those years and lost vision after the age of 10. We hypothesized that early blind people were expected to suffer the induced audiomotor conflict more than late blind and sighted, which in turn were expected not to differ.

The MAE analysis showed in every group significantly larger values when the auditory virtual reality platform altered rather than preserved the audiomotor loop. This result demonstrated our method's validity: the virtual platform altered the audiomotor loop in a behaviorally detectable way, yet it did not disrupt participants' ability to orient towards the target grossly. From the scientific standpoint, our method provided two results. First, the fact even sighted people were affected by the audiomotor conflict. In light of the well-established predictive coding theory, this is no surprise: since the brain works by predicting the perceptual outcome of motor actions [11], changing the perceptual consequences of motor commands was expected to affect spatial performance even in sighted people. The second outcome comes from the group-wise comparisons: larger MAE in early blind than in sighted people when the audiomotor loop was altered, but not when it was preserved. This result confirmed our hypothesis that early blind people rely on the audiomotor loop more than sighted people do, following both the theorized importance of vision for spatial calibration [2] and the hypothesized compensatory role of the audiomotor loop for spatial calibration in early blindness [4]. That said, our results did not identify any significant differences between the late blind group and the other ones. The late blind group's intermediate behavior contradicted our hypothesis that the compensatory role of the audiomotor loop is a prerogative of early blindness, suggesting its involvement in audiospatial processing is enhanced even in late blindness. This view is in line with previous findings that link modified audio-spatial information processing in late blindness and modifications in brain activity [12] and structure [13], mainly associated with the amount of time passed from blindness onset. Moreover, late blindness encompasses various conditions and progress, e.g., traumatic or degenerative, which are likely to rely differently on audiomotor information. Our late blind sample is not suitable for such further analysis because of its reduced size and scarce variability in disability histories. For these reasons, the link between types of late blindness and audiomotor loop use will be investigated in a future study.

### V. CONCLUSION

This work introduced a method to assess the reliance on the audiomotor coordination, or audiomotor loop, for audiospatial cues estimation using auditory virtual reality to alter the loop itself. Its use and utility were demonstrated in an experiment that compared the sensitivity to the audiomotor loop alteration as a function of blindness onset. A group of early blind, a group of late blind, and sighted people were tested on the platform. Results showed all the groups were sensitive to the induced audiomotor alteration, but early blind were significantly more so than sighted people; late blind people behaved intermediately. This study validated the platform efficacy and highlighted the increased reliance on the audiomotor loop for audio-spatial information processing in blindness, maybe even at late-onset. Such insights advocate for the development of new spatial orientation training for blind people based on exploiting the audiomotor loop, such as those developed by our group [14]. In the future, we will use our platform to address the relationship between the sensitivity to the audiomotor loop alteration and the years of blindness.

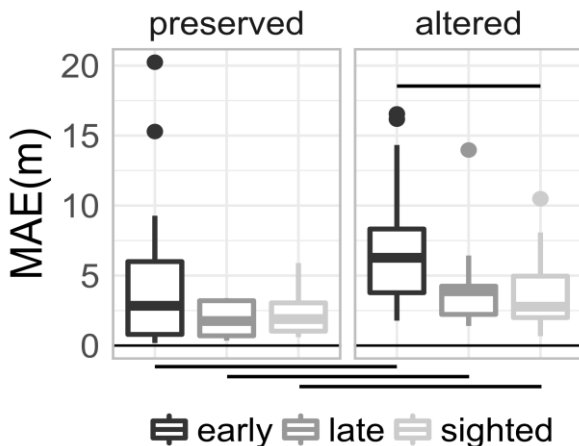


Fig. 4: MAE boxplots, split by age group and experimental condition. Boxes indicate 25,50 and 75 distribution percentiles. Outliers are depicted as full dots. Horizontal bars highlight significant post-hoc comparisons.

## ACKNOWLEDGEMENT

The research is partially supported by the MYSpace project to Monica Gori, which has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 948349).

## REFERENCES

- [1] M. Vasilyeva and S. F. Lourenco, "Development of spatial cognition," *Wiley Interdiscip. Rev. Cogn. Sci.*, vol. 3, no. 3, pp. 349–362, 2012.
- [2] M. Gori, G. Sandini, and D. Burr, "Development of Visuo-Auditory Integration in Space and Time," *Front. Integr. Neurosci.*, vol. 6, p. 77, Sep. 2012.
- [3] P. Voss, "Auditory Spatial Perception without Vision," *Front. Psychol.*, vol. 07, p. 1960, Dec. 2016.
- [4] J. Lewald, "Exceptional ability of blind humans to hear sound motion: Implications for the emergence of auditory space," *Neuropsychologia*, vol. 51, no. 1, pp. 181–186, Jan. 2013.
- [5] G. Virgili and G. Rubin, "Orientation and mobility training for adults with low vision," *Cochrane Database Syst. Rev.*, no. 5, May 2010.
- [6] S. Rohlf, L. Li, P. Bruns, and B. Röder, "Multisensory Integration Develops Prior to Crossmodal Recalibration," *Curr. Biol.*, vol. 30, no. 9, pp. 1726–1732.e7, May 2020.
- [7] Google, "Resonance Audio SDK for Unity." 2018.
- [8] ianovir, "HyperIMU." 2019.
- [9] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins, *The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only ANOVA Procedures*. 2011.
- [10] M. Gori, "Multisensory Integration and Calibration in Children and Adults with and without Sensory and Motor Disabilities," *Multisens. Res.*, vol. 28, no. 1–2, pp. 71–99, 2015.
- [11] Y. Huang and R. P. N. Rao, "Predictive coding," *Wiley Interdiscip. Rev. Cogn. Sci.*, vol. 2, no. 5, pp. 580–593, Sep. 2011.
- [12] A. Fieger, B. Röder, W. Teder-Sälejärvi, S. A. Hillyard, and H. J. Neville, "Auditory spatial tuning in late-onset blindness in humans," *J. Cogn. Neurosci.*, vol. 18, no. 2, pp. 149–157, Mar. 2006.
- [13] Q. Li, M. Song, J. Xu, W. Qin, C. Yu, and T. Jiang, "Cortical thickness development of human primary visual cortex related to the age of blindness onset," *Brain Imaging Behav.*, vol. 11, no. 4, pp. 1029–1036, Aug. 2017.
- [14] G. Cappagli *et al.*, "Audio motor training improves mobility and spatial cognition in visually impaired children," *Sci. Rep.*, vol. 9, no. 1, p. 3303, Dec. 2019.