Received: 23 January 2020 | Revised: 5 May 2020 | Accepted: 6 May 2020

DOI: 10.1111/ejn.14779

RESEARCH REPORT

EJN European Journal of Neuroscience FENS WILEY

Differential effects of the temporal and spatial distribution of audiovisual stimuli on cross-modal spatial recalibration

Patrick Bruns[1] | **Hubert R. Dinse**[2] | **Brigitte Röder**[1]

[1]Biological Psychology and Neuropsychology, University of Hamburg, Hamburg, Germany

[2]Neural Plasticity Lab, Institute of Neuroinformatics, Ruhr University Bochum, Bochum, Germany

Correspondence
Patrick Bruns, Biological Psychology and Neuropsychology, University of Hamburg, Von-Melle-Park 11, 20146 Hamburg, Germany.
Email: patrick.bruns@uni-hamburg.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: TRR 169/A1

Abstract

Visual input constantly recalibrates auditory spatial representations. Exposure to isochronous audiovisual stimuli with a fixed spatial disparity typically results in a subsequent auditory localization bias (ventriloquism aftereffect, VAE), whereas exposure to spatially congruent audiovisual stimuli improves subsequent auditory localization (multisensory enhancement, ME). Here, we tested whether cross-modal recalibration is affected by the stimulation rate and/or the distribution of audiovisual spatial disparities during training. Auditory localization was tested before and after participants were exposed either to audiovisual stimuli with a constant spatial disparity of 13.5° (VAE) or to spatially congruent audiovisual stimulation (ME). In a between-subjects design, audiovisual stimuli were presented either at a low frequency of 2 Hz, as used in previous studies of VAE and ME, or intermittently at a high frequency of 10 Hz, which mimics long-term potentiation (LTP) protocols and which was found superior in eliciting unisensory perceptual learning. Compared to low-frequency stimulation, VAE was reduced after high-frequency stimulation, whereas ME occurred regardless of the stimulation protocol. In two additional groups, we manipulated the spatial distribution of audiovisual stimuli in the low-frequency condition. Stimuli were presented with varying audiovisual disparities centered around 13.5° (VAE) or 0° (ME). Both VAE and ME were equally strong compared to a fixed spatial relationship of 13.5° or 0°, respectively. Taken together, our results suggest (a) that VAE and ME represent partly dissociable forms of learning and (b) that auditory representations adjust to the overall stimulus statistics rather than to a specific audiovisual spatial relationship.

KEYWORDS

long-term potentiation, multisensory, recalibration, spatial hearing, ventriloquism

Abbreviations: BF, Bayes factor; LTD, long-term depression; LTP, long-term potentiation; ME, multisensory enhancement; SC, superior colliculus; VAE, ventriloquism aftereffect.

Edited by John Foxe.

The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.14779

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2020 The Authors. *European Journal of Neuroscience* published by Federation of European Neuroscience Societies and John Wiley & Sons Ltd

1 | INTRODUCTION

Sensory maps are constantly calibrated by cross-modal input to achieve a coherent and accurate representation of the world. For example, after brief passive exposure to audiovisual stimuli with a consistent azimuthal spatial disparity, unisensory sound localization was found to be shifted in the direction of the spatially more reliable visual input, a phenomenon commonly referred to as the ventriloquism aftereffect (VAE) (Radeau & Bertelson, 1974; Recanzone, 1998). Similarly, sound localization performance was found to be improved after exposure to spatially aligned audiovisual stimuli (Passamonti, Frissen, & Làdavas, 2009; Strelnikov, Rosito, & Barone, 2011). Conversely, auditory stimuli were reported to facilitate the detection of both spatially aligned and spatially discrepant visual targets (Fiebelkorn, Foxe, Butler, & Molholm, 2011). Previous studies on the mechanisms underlying these cross-modal spatial recalibration effects have primarily tested the generalization of recalibration to untrained stimuli (e.g., sounds of different frequencies) or regions of space (Bertelson, Frissen, Vroomen, & de Gelder, 2006; Bruns & Röder, 2019b; Frissen, Vroomen, de Gelder, & Bertelson, 2003, 2005; Kopčo, Lin, Shinn-Cunningham, & Groh, 2009; Lewald, 2002; Recanzone, 1998; Sarlat, Warusfel, & Viaud-Delmon, 2006; Woods & Recanzone, 2004). However, little is known about the stimulation protocols that are optimal for inducing cross-modal recalibration effects in the first place.

Cross-modal recalibration emerges in parallel at two distinct time scales (Bosen, Fleming, Allen, O'Neill, & Paige, 2018; Bruns & Röder, 2015; Mendonça, Escher, van de Par, & Colonius, 2015; Van der Burg, Alais, & Cass, 2015; Watson, Akeroyd, Roach, & Webb, 2019). On the one hand, even a single exposure to a spatially discrepant audiovisual stimulus results in an immediate but transient adjustment of auditory localization (Wozny & Shams, 2011). On the other hand, repeated exposure to a consistent audiovisual disparity results in stronger and more stable changes in auditory processing that seem to integrate sensory evidence over time (Bruns & Röder, 2019a; Frissen, Vroomen, & de Gelder, 2012; Zierul, Röder, Tempelmann, Bruns, & Noesselt, 2017). To induce such cumulative cross-modal recalibration effects, previous studies have typically presented audiovisual stimuli with a fixed spatial relationship at a steady rate of one or two stimuli per second for several minutes (e.g., Bruns & Röder, 2019b; Frissen et al., 2012; Lewald, 2002; Passamonti et al., 2009; Radeau & Bertelson, 1974; Recanzone, 1998; Sarlat et al., 2006). This stimulation protocol reliably induced changes in sound localization behavior, but it remained unclear from these studies whether and how different temporal or spatial patterns of audiovisual stimulation would affect the strength of cross-modal recalibration. This knowledge would be informative for understanding the underlying neural mechanisms of cross-modal recalibration. Recently, we observed that exposure to incrementally increasing audiovisual discrepancies resulted in lower cross-modal recalibration than exposure to a constant audiovisual discrepancy, suggesting a crucial influence of the precise stimulation protocol on learning outcomes (Bruns & Röder, 2019a).

Exposure-based learning has most often been investigated in the context of unisensory perceptual learning studies (Beste & Dinse, 2013). For instance, prolonged passive sensory stimulation has been shown to enhance tactile acuity (Dinse, Ragert, Pleger, Schwenkreis, & Tegenthoff, 2003; Godde, Stauffenberg, Spengler, & Dinse, 2000), visual luminance change detection (Beste, Wascher, Güntürkün, & Dinse, 2011; Marzoll, Saygi, & Dinse, 2018), and face identity detection (Pegado, Vankrunkelsven, Steyaert, Boets, & Op de Beeck, 2016). Importantly, these unisensory learning effects depended on the temporal pattern of stimulation. Improvements were observed only when stimuli were presented intermittently at a high frequency of 10–20 Hz. By contrast, continuous stimulation at a low frequency of 1 Hz, which is similar to stimulation protocols commonly used to induce cross-modal recalibration, resulted in an impairment of visual and tactile discrimination performance (Beste et al., 2011; Pegado et al., 2016; Ragert, Kalisch, Bliem, Franzkowiak, & Dinse, 2008). In these studies, high- and low-frequency stimulation protocols were specifically designed to mimic protocols that are known to induce long-term potentiation (LTP) and long-term depression (LTD) of synaptic transmission at a cellular level (Bear & Malenka, 1994). Hence, it was concluded that high- and low-frequency sensory stimulation induced LTP- and LTD-like processes in the sensory systems, resulting in opposite effects on performance (Beste & Dinse, 2013; Dinse & Tegenthoff, 2018).

To the best of our knowledge, the effects of multisensory high-frequency stimulation have not been tested so far. However, cross-modal transfer of unisensory high-frequency stimulation effects has been demonstrated in a study in hemianopic patients with unilateral visual cortical lesions (Lewald, Tegenthoff, Peters, & Hausmann, 2012). In this study, high-frequency auditory stimulation in the blind hemifield of the patients induced a subsequent improvement of visual detection performance. At least in some cases, residual visual functions in hemianopic patients, known as blindsight, seem to depend on the functional integrity of the superior colliculus (SC) in the midbrain, which may convey visual information directly to extrastriate cortical areas (Fox, Goodale, & Bourne, 2020; Leopold, 2012). The SC is a major multisensory integration site that contains superimposed maps of auditory and visual space (Stein & Stanford, 2008), and might, thus, have mediated enhanced

visual performance in the hemianopic patients in response to auditory stimulation (Lewald et al., 2012).

Moreover, in other studies hemianopic patients showed an enhancement of auditory localization performance in response to (low-frequency) spatially congruent audiovisual stimulation in their blind hemifield (Leo, Bolognini, Passamonti, Stein, & Làdavas, 2008; Passamonti et al., 2009). In contrast, exposure to spatially discrepant audiovisual stimuli in the blind hemifield did not affect auditory localization performance. This dissociation between congruent and discrepant audiovisual stimulation suggests that flexible recalibration to changing cross-modal spatial correspondences requires visual cortical processing, whereas the collicular–extrastriate circuit allows for the integration of congruent audiovisual stimuli. Concordantly, studies in owls have suggested that spatial representations in the optic tectum, the avian homologue of the SC, become relatively stable in adulthood and are typically modulated by discrepant cross-modal input only within certain limits (Brainard & Knudsen, 1998).

Cross-modal recalibration studies have typically manipulated the spatial correspondence between auditory and visual stimuli without manipulating the variance of audiovisual discrepancies. One possibility is that presenting audiovisual stimuli with varying spatial discrepancies impedes recalibration because auditory and visual locations would be perceived as uncorrelated, and hence would be attributed to separate events (Parise, Spence, & Ernst, 2012). Another possibility is that auditory localization adjusts to the mean audiovisual discrepancy encountered in the exposure phase and hence would be unaffected by changes in variance. In the latter case, however, one might not expect an increase in auditory localization precision after exposure to varying audiovisual discrepancies with a mean of zero, because only spatially aligned audiovisual stimuli would activate the collicular–extrastriate pathway that appears to be crucial for reducing localization error (Leo et al., 2008; Passamonti et al., 2009). Concordantly, in studies with unimodal auditory stimulation it has been found that the perceived laterality of auditory stimuli was biased by the mean of recently encountered sound locations, whereas the sensitivity of lateralization judgments depended on the variance of the spatial distribution (Dahmen, Keating, Nodal, Schulz, & King, 2010).

In the present study, we systematically tested the effects of temporal (i.e., high- vs. low-frequency) and spatial (i.e., fixed vs. varying disparity) stimulation patterns on cross-modal recalibration to spatially discrepant (VAE) and spatially aligned (multisensory enhancement [ME]) audiovisual stimuli. In a between-subjects design, we compared recalibration effects in two baseline groups, which received standard exposure at a low frequency and with a fixed audiovisual disparity of either 13.5° or 0°, with two groups receiving exposure at a high frequency instead (Experiment 1) and two groups in which the audiovisual disparity varied around a mean of either 13.5° or 0° (Experiment 2). Two additional groups were exposed to auditory-only stimulation to control for unspecific test repetition effects (Experiment 3). We hypothesized (a) that high-frequency compared to low-frequency audiovisual stimulation would facilitate and (b) that exposure to varying audiovisual spatial discrepancies compared to a constant audiovisual correspondence would impede recalibration and audiovisual localization gains.

2 | MATERIALS AND METHODS

2.1 | Participants

A total of 120 healthy adult volunteers (28 male and 92 female) from the University of Hamburg community participated in the study. They had a mean age of 24.9 years (range 18–46 years), and all reported normal hearing and normal or corrected-to-normal vision. Participants were divided into eight groups of $n = 15$ each. All of them provided written informed consent prior to the study and received course credit or were compensated €7 for their participation. The experimental procedure was approved by the ethics commission of the German Psychological Society (DGPs), and the study was performed in accordance with the ethical standards laid down in the Declaration of Helsinki.

2.2 | Apparatus and stimuli

Participants were tested in a dark sound-attenuated room and faced the center of a semi-circular frame on which six loudspeakers (ConceptC Satellit, Teufel GmbH, Berlin, Germany) were mounted at a distance of 90 cm with eccentricities of ± 4.5°, ±13.5°, and ± 22.5° from the participants' straight-ahead position (0°). The loudspeakers were hidden from view behind an acoustically transparent curtain that extended to ± 90° from midline. For visual stimulation, a red laser pointer was projected onto the curtain at the level of the loudspeakers for 30 ms. The laser pointer was attached to a step motor which allowed visual stimulation at different azimuthal locations. Auditory stimuli were 1,000 Hz tones with a duration of 30 ms (including 5 ms linear rise/fall envelopes) that were presented at 65 dB(A). A sine tone, which is typically more difficult to localize than a noise burst, was used as in previous studies (e.g., Bruns & Röder, 2019a; Frissen et al., 2012; Lewald, 2002; Recanzone, 1998) to avoid potential ceiling effects in auditory localization performance. Sound intensity was randomly varied over a 4 dB range for every stimulus presentation to reduce any detectable differences in the loudspeaker transformation functions. During unimodal auditory localization tests, participants indicated

perceived sound locations with a rotatable hand pointer which was mounted in front of them.

2.3 | Pre- and posttest

All participants performed a unimodal sound localization task before (pretest) and immediately after (posttest) exposure to passive audiovisual (Experiments 1 and 2) or auditory-only (Experiment 3) stimulation. Pre- and posttests each consisted of 90 trials (15 trials per loudspeaker location) that were presented in a randomized order. At the beginning of each trial, the laser point was presented at the central azimuthal location (0°). The laser point was turned off as soon as the participants had aligned the hand pointer within ± 10° from center, and after a random delay (between 500 and 1,500 ms), an auditory target stimulus was presented from one of the six loudspeakers. Participants adjusted the hand pointer as accurately as possible toward the perceived azimuthal location of the sound source and confirmed their response by a button press. The next trial started 350 ms after the response.

2.4 | Exposure phase

After the auditory localization pretest, participants received one block of 600 exposure trials that had a duration of 5 min. The duration and the number of stimuli were identical for all experimental conditions. Auditory stimuli were always presented from the same six locations as in the pre- and posttest (100 stimuli per location). In Experiments 1 and 2, they were presented synchronously with a visual stimulus that was either spatially congruent or incongruent, depending on condition (see below). Stimuli were arranged in sets of 10 successive stimuli with identical locations. To ensure that participants attended to the stimulation, the laser point was occasionally illuminated in the time interval between the tenth stimulus of a set and the first stimulus of the next set. Participants had to press a button whenever they detected such a deviant stimulus, which occurred six times during the exposure phase. The pattern of sensory stimulation differed as follows between experiments and experimental conditions (see also Figure 1). Each participant completed one experiment and experimental condition only.

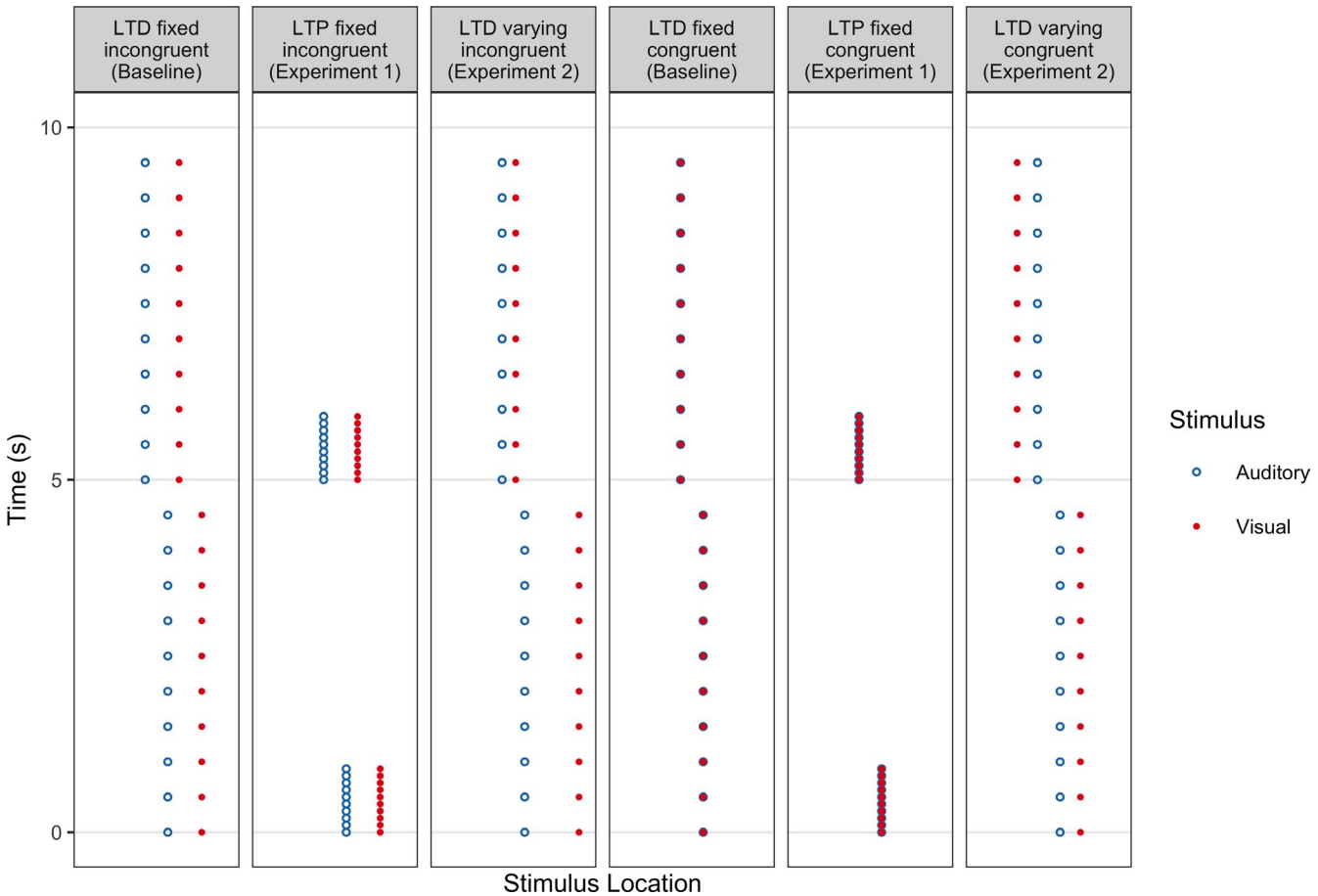

FIGURE 1 Schematic illustration of the audiovisual exposure phase. Stimulus locations changed every 5 s (i.e., after 10 audiovisual stimuli). Depending on condition, audiovisual stimuli were either presented at a steady rate of 2 Hz (LTD) or intermittently at 10 Hz (LTP). The audiovisual spatial relationship was either constant throughout the exposure phase (fixed) or changed every 5 s (varying). Visual stimuli were either (on average) displaced 13.5° to the right of the sound source (incongruent) or were (on average) presented at the same location as the sounds (congruent)

2.4.1 | Experiment 1

This experiment compared the effects of temporal high- versus low-frequency stimulation patterns on cross-modal recalibration, separately for spatially discrepant and spatially congruent audiovisual stimuli (see Figure 1). In the "LTD fixed incongruent" group, audiovisual stimuli were presented continuously at a low frequency of 2 Hz (i.e., two stimuli per second). Visual stimuli were always 13.5° to the right of the sound source. By contrast, in the "LTD fixed congruent" group, visual stimuli were always presented at the same location as the auditory stimuli (0° discrepancy). Aftereffects in the two LTD groups were compared with aftereffects in two groups which received high-frequency (LTP-like) stimulation instead. In the LTP groups, audiovisual stimuli were presented intermittently at a high frequency of 10 Hz. Sets of 10 stimuli (1 s stimulation) were separated by gaps of four seconds, amounting to 120 stimuli per minute as in the LTD conditions. In the "LTP fixed incongruent" group, visual stimuli were always 13.5° to the right of the sound source as in the "LTD fixed incongruent" group. In the "LTP fixed congruent" group, visual stimuli were always presented at the same location as the auditory stimuli (0° discrepancy) as in the "LTD fixed congruent" group.

2.4.2 | Experiment 2

This experiment tested the effects of a varying audiovisual spatial relationship on cross-modal recalibration in two new groups, as compared to the fixed relationship used in Experiment 1 (see Figure 1). In the "LTD varying incongruent" group, audiovisual stimuli were presented continuously at a low frequency of 2 Hz as in the LTD fixed conditions. However, the audiovisual spatial discrepancy varied between 5.4° and 21.6° (mean 13.5°) in steps of 1.8° with a uniform distribution. In the "LTD varying congruent" group, the audiovisual spatial discrepancies varied between −8.1° and 8.1° in steps of 1.8° with a mean of 0°; thus, although auditory and visual stimuli were never presented from congruent locations, their average locations were the same. Aftereffects in these two groups were compared to the two LTD groups from Experiment 1, in which the audiovisual spatial relationship was fixed at 13.5° ("LTD fixed incongruent" group) or 0° ("LTD fixed congruent" group).

2.4.3 | Experiment 3

To control for test repetition effects and unspecific effects of the stimulation protocols, two control groups were tested that received auditory-only stimulation during the exposure phase. In the first control group, auditory stimuli were presented continuously at a low frequency of 2 Hz ("LTD auditory control"). In the second control group, the auditory stimuli were presented intermittently at a high frequency of 10 Hz ("LTP auditory control").

2.5 | Data analysis

To verify that participants were able to perform the sound localization task at pretest, we calculated a simple linear regression between the mean pointing responses and the true locations of the six loudspeakers. Mean pointing responses were well approximated by regression lines in all participants ($R^2 \geq .66$, $p \leq .049$), except in one participant from the "LTP auditory control" group who was not able to discriminate the sound locations ($R^2 = 0.21$, $p = .367$). Data from this participant were, thus, excluded from further analyses.

Separate analyses were performed to assess VAEs in incongruent conditions and ME effects in congruent conditions (cf. Passamonti et al., 2009). VAEs in incongruent conditions were defined as rightward shifts in sound localization (i.e., in the direction of the visual stimuli) from pre- to posttest (Bertelson et al., 2006; Bruns & Röder, 2019b; Frissen, Vroomen, de Gelder, & Bertelson, 2003, 2005; Lewald, 2002; Recanzone, 1998). Thus, separately for the pre- and posttest, mean constant (i.e., signed) errors in sound localization were calculated by averaging responses across the 15 trials per loudspeaker location. Ventriloquism aftereffects were then calculated by subtracting constant errors in the pretest from those in the posttest and averaged across loudspeaker locations.

Multisensory enhancement effects in congruent conditions were defined as reductions in absolute localization errors (Passamonti et al., 2009; Strelnikov et al., 2011). For this purpose, mean absolute (i.e., unsigned) errors in sound localization were calculated by averaging absolute errors separately for the pre- and posttest. Posttest values were then subtracted from pretest values to yield positive values for error reductions after the exposure phase. The resulting values reflect both changes in localization accuracy and changes in localization precision. Potential ME effects could theoretically arise from the fact that the visual stimuli might have informed participants about the range of possible loudspeaker locations (−22.5° to 22.5°). To exclude that such higher-order knowledge accounted for performance improvements, we additionally tested for pre- to posttest changes in the slopes of the individually fitted regression lines between the mean pointing responses and the true locations of the six loudspeakers by means of paired *t* tests. A significant change in slope would indicate that the typically observed overshoot in localization responses with

pointing tasks (Bruns, Maiworm, & Röder, 2014; Bruns & Röder, 2019a; Lewald, 2002; Lewald & Getzmann, 2006; Zierul, Tong, Bruns, & Röder, 2019) was reduced by congruent audiovisual stimulation. Such a result would suggest that cognitive or decisional factors caused ME.

For comparisons of interest, independent two-sample *t* tests were used to test for differences in the size of the aftereffects between experimental conditions. Additionally, one-sample *t* tests were used to test whether aftereffects were significantly larger than zero. Where appropriate, *p* values were adjusted to control the family-wise error rate using Hochberg's step-up procedure of the Bonferroni correction (Hochberg, 1988). All *t* tests were additionally performed as Bayesian hypothesis tests in JASP version 0.12 (Wagenmakers et al., 2018), and Bayes Factors (BF_{10} for two-tailed tests and BF_{+0} for one-tailed tests) are reported. In addition, (Bayesian) one-way repeated-measures ANOVAs were used to test whether aftereffects significantly differed between the six loudspeaker locations. The Greenhouse-Geisser correction was used to correct for violations of the sphericity assumption and, where appropriate, the corrected *p* values are reported. Bayes Factors (BF_{incl}) for inclusion of the factor Location (as compared to the null model) are reported.

3 | RESULTS

3.1 | Experiment 1: Low- **versus** high-frequency stimulation

3.1.1 | Ventriloquism aftereffect

After audiovisual stimulation with a fixed spatial discrepancy of 13.5°, the "LTD fixed incongruent" group showed a clear shift in sound localization in the direction of the visual stimuli (i.e., a VAE). Compared to the LTD group, the VAE was significantly reduced in the "LTP fixed incongruent" group ($t_{28} = 2.30$, $p = .029$, $d = 0.84$; $BF_{10} = 2.32$; see Figure 2 left panel). Accordingly, one-sample *t* tests showed that the VAE was significantly larger than zero in the "LTD fixed incongruent" group ($M = 4.7°$, $SEM = 0.7$; $t_{14} = 6.25$, $p < .001$, $d = 1.61$; $BF_{+0} = 2{,}244.34$), but not in the "LTP fixed incongruent" group ($M = 1.7°$, $SEM = 1.1$; $t_{14} = 1.63$, $p = .062$, $d = 0.42$; $BF_{+0} = 1.44$). The group-averaged sound localization responses per loudspeaker location are shown in Figure 3. As can be seen, the VAE in the "LTD fixed incongruent" group did not differ between loudspeaker locations ($F_{5,70} = 0.30$, $p = .909$, $\eta^2_G = 0.02$; $BF_{incl} = 0.06$).

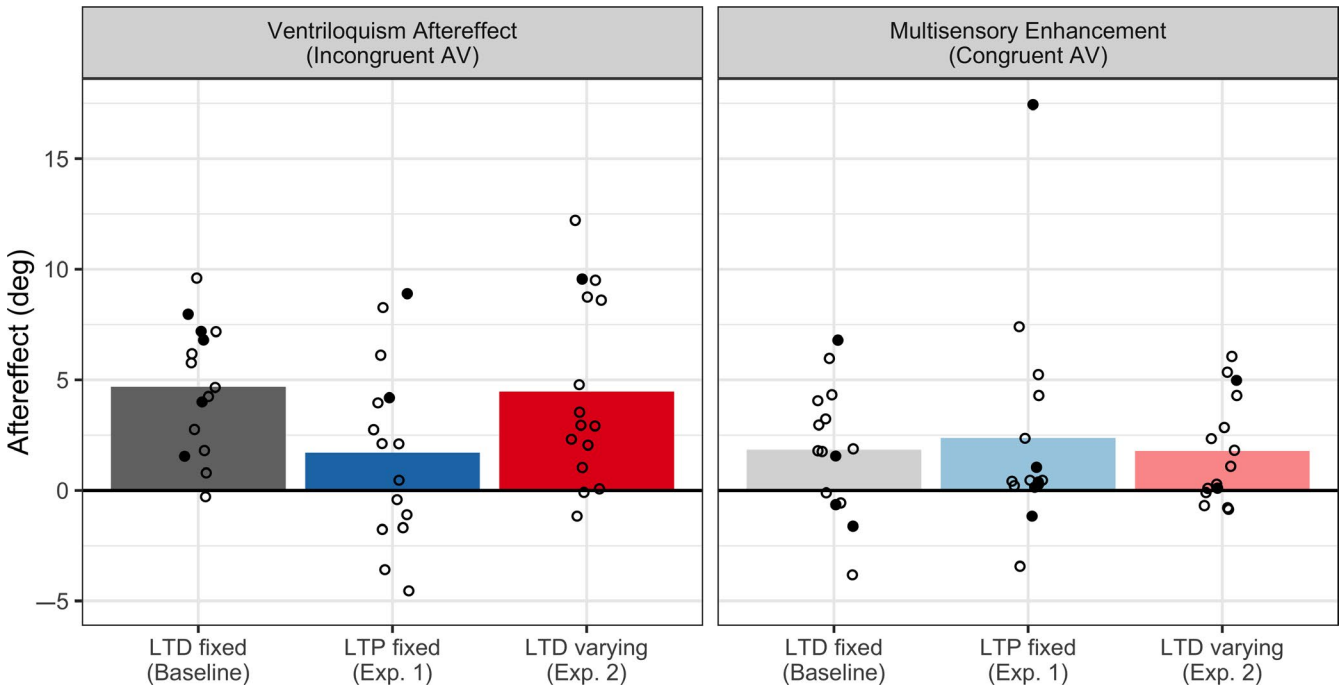

FIGURE 2 Aftereffects in Experiments 1 and 2. Left panel shows mean rightward shifts in sound localization (i.e., ventriloquism aftereffects) at posttest (as compared to pretest) in conditions with incongruent audiovisual exposure. Right panel shows mean reductions in absolute localization errors (i.e., multisensory enhancements) at posttest (as compared to pretest) in conditions with congruent audiovisual exposure. Both panels use the same *y*-axis scale. Bars indicate group means and superimposed dots indicate individual means (white dots indicate female participants and black dots indicate male participants)

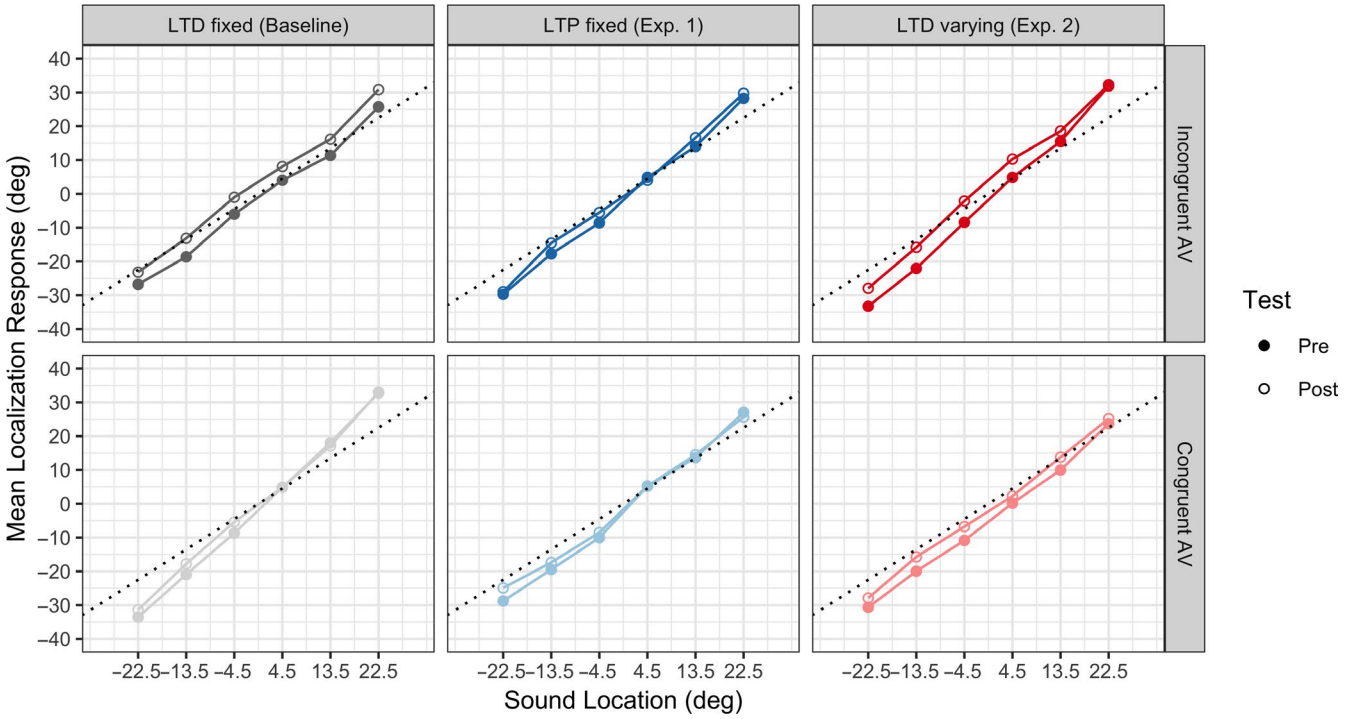

FIGURE 3 Group-averaged sound localization responses per loudspeaker location at pretest and at posttest in Experiments 1 and 2. Dots indicate mean pointing responses (in degrees) across the 15 trials per loudspeaker location, averaged across participants in each group. Dotted lines indicate the actual locations of the loudspeakers. Differences between pretest and posttest in incongruent audiovisual conditions correspond to ventriloquism aftereffects. Differences between pretest and posttest in congruent audiovisual conditions reflect changes in accuracy only

3.1.2 | Multisensory enhancement

Both "fixed congruent" groups showed a reduction in absolute sound localization errors after exposure to spatially congruent audiovisual stimulation. The size of the aftereffect did not differ significantly between the two groups ($t_{28} = 0.36$, $p = .721$, $d = 0.13$; $BF_{10} = 0.36$; see Figure 2 right panel). Accordingly, the localization error reduction was significantly larger than zero both in the "LTD fixed congruent" group ($M = 1.8°$, $SEM = 0.7$; $t_{14} = 2.47$, $p = .026$, $d = 0.64$; $BF_{+0} = 4.92$) and the "LTP fixed congruent" group ($M = 2.4°$, $SEM = 1.3$; $t_{14} = 1.86$, $p = .042$, $d = 0.48$; $BF_{+0} = 1.96$). Localization error reductions did not significantly differ between locations, neither in the "LTD fixed congruent" group ($F_{5,70} = 1.93$, $p = .149$, $\eta^2_G = 0.07$; $BF_{incl} = 0.65$) nor in the "LTP fixed congruent" group ($F_{5,70} = 2.09$, $p = .077$, $\eta^2_G = 0.04$; $BF_{incl} = 0.74$), suggesting that error reductions reflected location-unspecific improvements in accuracy and precision. Accordingly, localization accuracy did not selectively improve for peripheral locations (see Figure 3): There were no significant changes in the slopes of the individually fitted regression lines, neither in the "LTD fixed congruent" group ($t_{14} = 1.84$, $p = .088$, $d = 0.47$; $BF_{10} = 1.00$) nor in the "LTP fixed congruent" group ($t_{14} = 1.28$, $p = .223$, $d = 0.33$; $BF_{10} = 0.52$).

3.2 | Experiment 2: Varying versus fixed spatial relationship

3.2.1 | Ventriloquism aftereffect

After stimulation with varying audiovisual discrepancies centered around 13.5°, the "LTD varying incongruent" group showed a significant VAE ($M = 4.5°$, $SEM = 1.1$; $t_{14} = 4.12$, $p < .001$, $d = 1.06$; $BF_{+0} = 74.90$). The VAE differed between locations ($F_{5,70} = 3.70$, $p = .017$, $\eta^2_G = 0.12$; $BF_{incl} = 8.77$) but was significant at all locations ($ts \geq 2.61$, $ps \leq 0.02$, $ds \geq 0.67$; $BF_{+0} \geq 6.11$) except at the rightmost location ($t_{14} = 0.24$, $p = .408$, $d = 0.06$; $BF_{+0} = 0.32$; see Figure 3). Moreover, the overall size of the VAE did not differ significantly from the "LTD fixed incongruent" group in Experiment 1 which received audiovisual stimulation with a fixed spatial discrepancy of 13.5° ($t_{28} = 0.16$, $p = .872$, $d = 0.06$; $BF_{10} = 0.35$; see Figure 2 left panel).

3.2.2 | Multisensory enhancement

The "LTD varying congruent" group showed a significant reduction of absolute sound localization errors after exposure to audiovisual stimuli with varying discrepancies centered around 0° ($M = 1.8°$, $SEM = 0.6$; $t_{14} = 2.88$, $p = .006$,

d = 0.74; BF_{+0} = 9.48). The localization error reduction did not significantly differ between locations ($F_{5,70}$ = 1.03, p = .390, η^2_G = 0.04; BF_{incl} = 0.16), and there was no significant change in the slope of the individually fitted regression lines (t_{14} = 0.76, p = .459, d = 0.20; BF_{10} = 0.34). The overall size of the error reduction did not differ significantly from the "LTD fixed congruent" group in Experiment 1 in which audiovisual stimuli were always spatially congruent (t_{28} = 0.05, p = .959, d = 0.02; BF_{10} = 0.35; see Figure 2 right panel).

3.3 | Experiment 3: Auditory-only stimulation

3.3.1 | Localization bias

Compared to pretest, sound localization responses in both control groups were numerically shifted toward the right after auditory-only stimulation (see Figure 4 left panel). However, these localization biases did not differ significantly from zero, neither in the "LTD auditory control" group (M = 0.8°, SEM = 0.9; t_{14} = 0.90, p = .382, d = 0.23; BF_{10} = 0.37), nor in the "LTP auditory control" group (M = 2.5°, SEM = 1.5; t_{13} = 1.67, p = .119, d = 0.45; BF_{10} = 0.82). Moreover, the localization bias in the "LTD auditory control" group was significantly smaller than the VAEs observed in the "LTD fixed incongruent" group from Experiment 1 (t_{28} = 3.31, p = .003, d = 1.21; BF_{+0} = 29.10) and the "LTD varying incongruent" group from Experiment 2 (t_{28} = 2.60, p = .014, d = 0.95; BF_{+0} = 7.52). By contrast, the localization bias in the "LTP auditory control" group did not differ significantly from the "LTP fixed incongruent" group (t_{27} = 0.45, p = .671, d = 0.17; BF_{+0} = 0.26).

3.3.2 | Absolute localization error

Numerically, both control groups showed an increase rather than a decrease in absolute localization errors after auditory-only stimulation, as compared to the pretest (see Figure 4 right panel). These changes in localization error did not differ significantly from zero, neither in the "LTD auditory control" group (M = −1.3°, SEM = 1.5; t_{14} = 0.87, p = .398, d = 0.23; BF_{10} = 0.37), nor in the "LTP auditory control" group (M = −2.3°, SEM = 1.8; t_{13} = 1.26, p = .232, d = 0.34; BF_{10} = 0.52). Moreover, error reductions observed in the three experimental groups from Experiments 1 and 2 were significantly larger compared to their respective control group ("LTD fixed congruent" vs "LTD auditory control": t_{28} = 1.87, p = .036, d = 0.68; BF_{+0} = 2.35; "LTP fixed congruent" vs. "LTP auditory control": t_{27} = 2.12, p = .036, d = 0.79; BF_{+0} = 3.38; "LTD varying congruent" vs. "LTD auditory control": t_{28} = 1.90, p = .036, d = 0.69; BF_{+0} = 2.45).

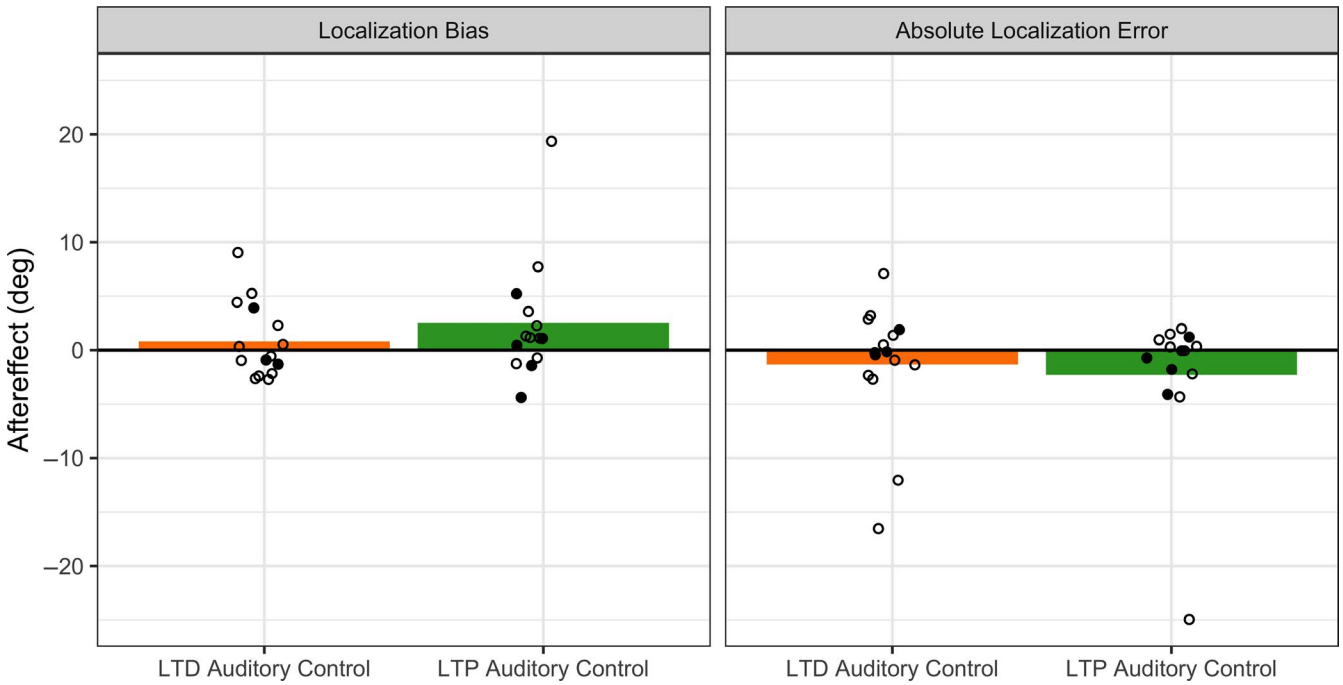

FIGURE 4 Changes in sound localization performance after auditory-only stimulation in Experiment 3. Left panel shows mean shifts in sound localization (i.e., localization bias) at posttest as compared to pretest (positive values indicate rightward shifts and negative values indicate leftward shifts). Right panel shows changes in absolute localization error at posttest as compared to pretest (positive values indicate error reductions and negative values indicate error increases). Both panels use the same y-axis scale. Bars indicate group means, and superimposed dots indicate individual means (white dots indicate female participants, and black dots indicate male participants)

Note that one participant in the "LTP fixed congruent" group showed an unusually large ME effect of 17.4°, whereas one participant in the "LTP auditory control" group showed a large decrease of 25.0° in performance (these two were the only values that were more than 3 *SD*s from their respective group means). However, the difference between the two conditions remained significant after exclusion of these two participants ($t_{25} = 1.96$, $p = .036$, $d = 0.76$; $BF_{+0} = 2.68$).

4 | DISCUSSION

In order to shed light on the neural mechanisms of cross-modal learning, we examined how the temporal pattern of audiovisual stimulation (Experiment 1) and the distribution of audiovisual spatial discrepancies (Experiment 2) affect the strength of cross-modal spatial recalibration. These manipulations additionally allowed us to compare different types of cross-modal learning (VAE and ME) with unisensory perceptual learning for which different temporal stimulation patterns resulted in opposite effects (i.e., improvements vs. impairments) on performance. As in previous studies, continuous audiovisual stimulation at a low frequency of 2 Hz reliably induced a subsequent shift in sound localization toward the visual attractor (i.e., a VAE) in the incongruent condition (13.5° audiovisual disparity) and reduced sound localization errors in the congruent condition (0° audiovisual disparity). However, changing the temporal stimulation pattern to an intermittent presentation of the audiovisual stimuli at a high frequency of 10 Hz significantly reduced the VAE, although it did not affect ME after spatially congruent audiovisual stimulation (Experiment 1). By contrast, the distribution of audiovisual spatial discrepancies (either always the same or varying between trials) had no effect on cross-modal recalibration, as both VAEs and ME effects did not differ between the two conditions (Experiment 2). A control experiment confirmed that both aftereffects were due to cross-modal stimulation during the exposure phase and did not occur after auditory-only stimulation (Experiment 3).

Why did the temporal pattern of audiovisual stimulation (low- vs. high-frequency) selectively modulate the VAE but not ME? Previous research has indicated that these two types of cross-modal recalibration are mediated by dissociable mechanisms. Recalibration to an altered audiovisual spatial correspondence in the VAE required intact visual cortical processing, whereas ME after congruent audiovisual stimulation did not and rather seemed to involve a collicular–extrastriate circuit (Bertini, Leo, Avenanti, & Làdavas, 2010; Leo et al., 2008; Passamonti et al., 2009). Our results provide corroborating evidence for this dissociation and might suggest that high-frequency stimulation made a difference only for cortical processing underlying the VAE.

Previous event-related potential studies suggested that it takes the brain about 160 ms to detect an audiovisual spatial discrepancy, likely in multisensory parietal regions (Gondan, Niederhaus, Rösler, & Röder, 2005; Teder-Sälejärvi, Di Russo, McDonald, & Hillyard, 2005). Parietal regions might, however, not be necessary for integrating auditory and visual spatial information (Phan, Schendel, Recanzone, & Robertson, 2000). Rather, audiovisual spatial information might be looped back from parietal regions to auditory cortex, resulting in an altered auditory activation around 260 ms post-stimulus which mediates the behavioral ventriloquism effect (Bonath et al., 2007; Bruns & Röder, 2010; Stekelenburg, Vroomen, & de Gelder, 2004). Recently, it has been shown that the VAE depends on the same neural pathway (Park & Kayser, 2019; Zierul et al., 2017), suggesting that a stimulation frequency of less than 4 Hz (i.e., 250 ms stimulus onset asynchrony) might be necessary for inducing a cumulative VAE.

An upper limit of around 4 Hz for successful task performance had previously been reported for tasks in which participants had to judge audiovisual temporal synchrony or had to match individual visual and auditory stimuli—at stimulation rates above 4 Hz participants were no longer able to perform these tasks above chance (Benjamins, van der Smagt, & Verstraten, 2008; Fujisaki & Nishida, 2005, 2007; for review, see Chen & Vroomen, 2013). These findings do, however, not necessarily imply that the individual stimuli were not perceived as separate events. In fact, for synchronous audiovisual stimulation (as used in the present study), it has been demonstrated that participants are able to discriminate stimulation rates around 10 Hz and that they benefit from audiovisual as compared to unimodal stimulation which is consistent with statistically optimal multisensory integration (Brooks, Anderson, Roach, McGraw, & McKendrick, 2015; Locke & Landy, 2017). Thus, high-frequency stimulation does not seem to interfere with genuine multisensory integration. Rather, specific multisensory functions as judging the temporal (or spatial) alignment might be impaired by high-frequency stimulation, in line with the present results which found an effect of stimulation rate only for the VAE but not for the ME effect.

One possibility is that subcortical processes, which are assumed to mediate the ME effect, are not responsive to stimulation frequency. If ME in response to congruent audiovisual stimulation is predominantly due to changes in the receptive fields of audiovisual neurons in the superficial layers of the SC, such a calibration would not require a time-consuming looping through distant neural circuits and might just depend on the co-activation statistics, rendering ME independent of the stimulation frequency. Alternatively, the audiovisual exposure duration of only 5 min in the present study might have been too short to show a differential effectiveness of high- versus low-frequency stimulation for the ME effect.

Earlier studies showed that longer unimodal high-frequency stimulation for one hour altered visual collicular–extrastriate processing in hemianopic patients (Lewald et al., 2012). In any case, the presence of a ME effect with high-frequency stimulation suggests that the selective reduction of the VAE was due to a specific temporal limitation of the neural circuitry required for cross-modal recalibration and not due to a general impairment of multisensory integration or other unspecific effects of high-frequency stimulation.

Our finding that the VAE was reduced and not significant with high-frequency audiovisual stimulation might seem counterintuitive, given that unisensory perceptual learning studies have typically reported enhanced performance following high-frequency stimulation (Beste et al., 2011; Klein, Magerl, Hopf, Sandkühler, & Treede, 2004; Marzoll et al., 2018; Pegado et al., 2016; Ragert et al., 2008). However, at least in tactile perception both improvements and impairments have been observed as a result of high-frequency stimulation. For example, improvements in tactile acuity (i.e., discriminating two closely spaced stimuli) developed at the cost of localization performance (i.e., accurately telling the stimulated skin location), implying a trade-off between discrimination and localization (Dinse, Kattenstroh, Gatica Tossi, Tegenthoff, & Kalisch, 2011; Dinse, Wilimzig, & Kalisch, 2008).

Modeling approaches have shown that spatial discrimination and localization are oppositely affected by lateral inhibition in a spatiotopically organized population of neurons, as typically found in somatosensory and visual cortex (Dinse & Jancke, 2001; Dinse et al., 2008; Jancke et al., 1999). On the one hand, strong lateral inhibition impairs tactile two-point discrimination performance because it reduces coexisting activation at other locations in the cortical population of neurons, which is necessary for discriminating two points from one. On the other hand, lateral inhibition reduces noisy fluctuations of activation, thereby improving localization of the peak position in the cortical map. The model assumes that high-frequency tactile stimulation induces a decrease of lateral inhibition, which would result in enhanced two-point discrimination performance due to stronger bimodal activations when two stimuli are present, but would impair localization performance due to increased noise in determining the location of a single activation peak (Dinse et al., 2008).

In contrast to the previous findings in vision and touch (Dinse et al., 2008; Dinse et al., 2011; Jancke et al., 1999), high-frequency auditory-only stimulation in the present study did not significantly impair auditory localization performance. Auditory spatial perception might be less influenced by a potential decrease in lateral inhibition after high-frequency stimulation, because neurons in auditory cortex do not form a spatiotopic map as found in the visual and somatosensory systems (McAlpine & Grothe, 2003; Stecker, Harrington, & Middlebrooks, 2005). We did not directly test the effect of high-frequency visual stimulation on visual localization performance. However, if high-frequency stimulation indeed impaired visual spatial reliability to a larger extent than auditory spatial reliability, this could explain why the VAE decreased after audiovisual high-frequency stimulation. Reducing the relative reliability of the visual stimuli has been shown to result in a reduced visual capture of the perceived auditory location (Alais & Burr, 2004), which in turn could have resulted in a smaller cross-modal recalibration effect. However, some studies have reported that the strength of cross-modal recalibration does not depend on the relative reliability of the cross-modal stimuli (Zaidel, Turner, & Angelaki, 2011).

In addition, or as an alternative to a modality-specific effect on visual spatial reliability, high-frequency stimulation might have reduced the reliability of the audiovisual spatial disparity itself. If processing of the size of the disparity between the visual and the auditory stimulus became more variable under high-frequency stimulation, cross-modal recalibration might have been inhibited because the stimulation would be less informative about the true audiovisual spatial correspondence. However, the results of Experiment 2 largely rule out this possibility. In Experiment 2, audiovisual stimuli were presented with varying disparities, a manipulation which should have resembled the effect of a reduced reliability of audiovisual spatial processing. Nevertheless, the size of the aftereffects was virtually identical to those seen after audiovisual stimulation with a fixed disparity.

The finding that the distribution of audiovisual spatial discrepancies (fixed vs. varying) did not affect the strength of cross-modal recalibration might suggest that sensory evidence was integrated over time and auditory spatial representations were adapted by the mean audiovisual disparity rather than on a trial-by-trial basis. This could explain why sound localization improved in the varying congruent condition (mean disparity of 0°) although audiovisual stimuli were never actually presented from the same location in this condition. Previous studies have shown that tones of two different sound frequencies can be recalibrated independently even if they were concurrently paired with visual stimuli that had a mean disparity (across sound frequencies) of zero (Bruns & Röder, 2015, 2019b). Since this effect was dissociable from a trial-by-trial influence on unisensory sound localization (Bruns & Röder, 2015), it might similarly indicate an integration of sensory evidence over time. In the present study, the presentation of audiovisual exposure trials and unimodal auditory test trials in separate blocks might have particularly facilitated integration of sensory information over trials, because the task did not require any localization responses during the audiovisual exposure phase.

In summary, our results substantiate previous evidence (Passamonti et al., 2009) indicating that spatially congruent and spatially discrepant audiovisual stimulation activate dissociable learning mechanisms and show that they are

differently sensitive to the temporal pattern of audiovisual stimulation. By contrast, changes in the variability of audiovisual discrepancies did not affect learning, suggesting that sensory representations adjust to the overall stimulus statistics rather than to a specific cross-modal relationship. These findings have implications for the design of multisensory training protocols which should carefully consider potential effects of the chosen temporal and spatial pattern of cross-modal stimulation on learning outcomes.

ACKNOWLEDGEMENTS

This work was supported by German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) Grant TRR 169/A1. We thank Elisa Bußkamp, Lisa Dräger, and Samantha Schröder for help collecting data.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

AUTHOR CONTRIBUTIONS

All authors conceptualized the study, interpreted the results, and reviewed the final manuscript. Patrick Bruns performed the experiments, analyzed the data, and wrote the manuscript.

DATA AVAILABILITY STATEMENT

The raw data associated with this study are publicly available in the UHH Research Data Repository at http://doi.org/10.25592/uhhfdm.948

ORCID

Patrick Bruns https://orcid.org/0000-0002-2977-8874

REFERENCES

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*, 257–262. https://doi.org/10.1016/j.cub.2004.01.029

Bear, M. F., & Malenka, R. C. (1994). Synaptic plasticity: LTP and LTD. *Current Opinion in Neurobiology*, *4*, 389–399. https://doi.org/10.1016/0959-4388(94)90101-5

Benjamins, J. S., van der Smagt, M. J., & Verstraten, F. A. J. (2008). Matching auditory and visual signals: Is sensory modality just another feature? *Perception*, *37*, 848–858. https://doi.org/10.1068/p5783

Bertelson, P., Frissen, I., Vroomen, J., & de Gelder, B. (2006). The aftereffects of ventriloquism: Patterns of spatial generalization. *Perception & Psychophysics*, *68*, 428–436. https://doi.org/10.3758/BF03193687

Bertini, C., Leo, F., Avenanti, A., & Làdavas, E. (2010). Independent mechanisms for ventriloquism and multisensory integration as revealed by theta-burst stimulation. *European Journal of Neuroscience*, *31*, 1791–1799. https://doi.org/10.1111/j.1460-9568.2010.07200.x

Beste, C., & Dinse, H. R. (2013). Learning without training. *Current Biology*, *23*, R489–R499. https://doi.org/10.1016/j.cub.2013.04.044

Beste, C., Wascher, E., Güntürkün, O., & Dinse, H. R. (2011). Improvement and impairment of visually guided behavior through LTP- and LTD-like exposure-based visual learning. *Current Biology*, *21*, 876–882. https://doi.org/10.1016/j.cub.2011.03.065

Bonath, B., Noesselt, T., Martinez, A., Mishra, J., Schwiecker, K., Heinze, H.-J., & Hillyard, S. A. (2007). Neural basis of the ventriloquist illusion. *Current Biology*, *17*, 1697–1703. https://doi.org/10.1016/j.cub.2007.08.050

Bosen, A. K., Fleming, J. T., Allen, P. D., O'Neill, W. E., & Paige, G. D. (2018). Multiple time scales of the ventriloquism aftereffect. *PLoS One*, *13*, e0200930. https://doi.org/10.1371/journal.pone.0200930

Brainard, M. S., & Knudsen, E. I. (1998). Sensitive periods for visual calibration of the auditory space map in the barn owl optic tectum. *Journal of Neuroscience*, *18*, 3929–3942. https://doi.org/10.1523/JNEUROSCI.18-10-03929.1998

Brooks, C. J., Anderson, A. J., Roach, N. W., McGraw, P. V., & McKendrick, A. M. (2015). Age-related changes in auditory and visual interactions in temporal rate perception. *Journal of Vision*, *15*, 2. https://doi.org/10.1167/15.16.2

Bruns, P., Maiworm, M., & Röder, B. (2014). Reward expectation influences audiovisual spatial integration. *Attention, Perception, & Psychophysics*, *76*, 1815–1827. https://doi.org/10.3758/s13414-014-0699-y

Bruns, P., & Röder, B. (2010). Tactile capture of auditory localization: An event-related potential study. *European Journal of Neuroscience*, *31*, 1844–1857. https://doi.org/10.1111/j.1460-9568.2010.07232.x

Bruns, P., & Röder, B. (2015). Sensory recalibration integrates information from the immediate and the cumulative past. *Scientific Reports*, *5*, 12739. https://doi.org/10.1038/srep12739

Bruns, P., & Röder, B. (2019a). Repeated but not incremental training enhances cross-modal recalibration. *Journal of Experimental Psychology: Human Perception and Performance*, *45*, 435–440. https://doi.org/10.1037/xhp0000642

Bruns, P., & Röder, B. (2019b). Spatial and frequency specificity of the ventriloquism aftereffect revisited. *Psychological Research Psychologische Forschung*, *83*, 1400–1415. https://doi.org/10.1007/s00426-017-0965-4

Chen, L., & Vroomen, J. (2013). Intersensory binding across space and time: A tutorial review. *Attention, Perception, & Psychophysics*, *75*, 790–811. https://doi.org/10.3758/s13414-013-0475-4

Dahmen, J. C., Keating, P., Nodal, F. R., Schulz, A. L., & King, A. J. (2010). Adaptation to stimulus statistics in the perception and neural representation of auditory space. *Neuron*, *66*, 937–948. https://doi.org/10.1016/j.neuron.2010.05.018

Dinse, H. R., & Jancke, D. (2001). Comparative population analysis of cortical representations in parametric spaces of visual field and skin: A unifying role for nonlinear interactions as a basis for active information processing across modalities. *Progress in Brain Research*, *130*, 155–173. https://doi.org/10.1016/S0079-6123(01)30011-0

Dinse, H. R., Kattenstroh, J. C., Gatica Tossi, M. A., Tegenthoff, M., & Kalisch, T. (2011). Sensory stimulation for augmenting perception, sensorimotor behavior and cognition. In I. Segev, & H. Markram (Eds.), *Augmenting cognition* (pp. 11–39). Lausanne: EPFL Press.

Dinse, H. R., Ragert, P., Pleger, B., Schwenkreis, P., & Tegenthoff, M. (2003). Pharmacological modulation of perceptual learning and associated cortical reorganization. *Science*, *301*, 91–94. https://doi.org/10.1126/science.1085423

Dinse, H. R., & Tegenthoff, M. (2018). Repetitive sensory stimulation – A canonical approach to control the induction of human learning at a

behavioral and neural level. *Handbook of Behavioral Neuroscience*, *28*, 389–413. https://doi.org/10.1016/B978-0-12-812028-6.00021-5

Dinse, H. R., Wilimzig, C., & Kalisch, T. (2008). Learning effects in haptic perception. In M. Grunwald (Ed.), *Human haptic perception: Basics and applications* (pp. 165–182). Basel: Birkhäuser Verlag.

Fiebelkorn, I. C., Foxe, J. J., Butler, J. S., & Molholm, S. (2011). Auditory facilitation of visual-target detection persists regardless of retinal eccentricity and despite wide audiovisual misalignments. *Experimental Brain Research*, *213*, 167–174. https://doi.org/10.1007/s00221-011-2670-7

Fox, D. M., Goodale, M. A., & Bourne, J. A. (2020). The age-dependent neural substrates of blindsight. *Trends in Neurosciences*, *43*, 242–252. https://doi.org/10.1016/j.tins.2020.01.007

Frissen, I., Vroomen, J., & de Gelder, B. (2012). The aftereffects of ventriloquism: The time course of the visual recalibration of auditory localization. *Seeing and Perceiving*, *25*, 1–14. https://doi.org/10.1163/187847611X620883

Frissen, I., Vroomen, J., de Gelder, B., & Bertelson, P. (2003). The aftereffects of ventriloquism: Are they sound-frequency specific? *Acta Psychologica*, *113*, 315–327. https://doi.org/10.1016/S0001-6918(03)00043-X

Frissen, I., Vroomen, J., de Gelder, B., & Bertelson, P. (2005). The aftereffects of ventriloquism: Generalization across sound-frequencies. *Acta Psychologica*, *118*, 93–100. https://doi.org/10.1016/j.actpsy.2004.10.004

Fujisaki, W., & Nishida, S. (2005). Temporal frequency characteristics of synchrony-asynchrony discrimination of audio-visual signals. *Experimental Brain Research*, *166*, 455–464. https://doi.org/10.1007/s00221-005-2385-8

Fujisaki, W., & Nishida, S. (2007). Feature-based processing of audio-visual synchrony perception revealed by random pulse trains. *Vision Research*, *47*, 1075–1093. https://doi.org/10.1016/j.visres.2007.01.021

Godde, B., Stauffenberg, B., Spengler, F., & Dinse, H. R. (2000). Tactile coactivation-induced changes in spatial discrimination performance. *Journal of Neuroscience*, *20*, 1597–1604. https://doi.org/10.1523/JNEUROSCI.20-04-01597.2000

Gondan, M., Niederhaus, B., Rösler, F., & Röder, B. (2005). Multisensory processing in the redundant-target effect: A behavioral and event-related potential study. *Perception & Psychophysics*, *67*, 713–726. https://doi.org/10.3758/BF03193527

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, *75*, 800–802. https://doi.org/10.1093/biomet/75.4.800

Jancke, D., Erlhagen, W., Dinse, H. R., Akhavan, A. C., Giese, M., Steinhage, A., & Schöner, G. (1999). Parametric population representation of retinal location: Neuronal interaction dynamics in cat primary visual cortex. *Journal of Neuroscience*, *19*, 9016–9028. https://doi.org/10.1523/JNEUROSCI.19-20-09016.1999

Klein, T., Magerl, W., Hopf, H.-C., Sandkühler, J., & Treede, R.-D. (2004). Perceptual correlates of nociceptive long-term potentiation and long-term depression in humans. *Journal of Neuroscience*, *24*, 964–971. https://doi.org/10.1523/JNEUROSCI.1222-03.2004

Kopčo, N., Lin, I.-F., Shinn-Cunningham, B. G., & Groh, J. M. (2009). Reference frame of the ventriloquism aftereffect. *Journal of Neuroscience*, *29*, 13809–13814. https://doi.org/10.1523/JNEUROSCI.2783-09.2009

Leo, F., Bolognini, N., Passamonti, C., Stein, B. E., & Làdavas, E. (2008). Cross-modal localization in hemianopia: New insights on multisensory integration. *Brain*, *131*, 855–865. https://doi.org/10.1093/brain/awn003

Leopold, D. A. (2012). Primary visual cortex: Awareness and blindsight. *Annual Review of Neuroscience*, *35*, 91–109. https://doi.org/10.1146/annurev-neuro-062111-150356

Lewald, J. (2002). Rapid adaptation to auditory-visual spatial disparity. *Learning & Memory*, *9*, 268–278. https://doi.org/10.1101/lm.51402

Lewald, J., & Getzmann, S. (2006). Horizontal and vertical effects of eye-position on sound localization. *Hearing Research*, *213*, 99–106. https://doi.org/10.1016/j.heares.2006.01.001

Lewald, J., Tegenthoff, M., Peters, S., & Hausmann, M. (2012). Passive auditory stimulation improves vision in hemianopia. *PLoS One*, *7*, e31603. https://doi.org/10.1371/journal.pone.0031603

Locke, S. M., & Landy, M. S. (2017). Temporal causal inference with stochastic audiovisual sequences. *PLoS One*, *12*, e0183776. https://doi.org/10.1371/journal.pone.0183776

Marzoll, A., Saygi, T., & Dinse, H. R. (2018). The effect of LTP- and LTD-like visual stimulation on modulation of human orientation discrimination. *Scientific Reports*, *8*, 16156. https://doi.org/10.1038/s41598-018-34276-z

McAlpine, D., & Grothe, B. (2003). Sound localization and delay lines – Do mammals fit the model? *Trends in Neurosciences*, *26*, 347–350. https://doi.org/10.1016/S0166-2236(03)00140-1

Mendonça, C., Escher, A., van de Par, S., & Colonius, H. (2015). Predicting auditory space calibration from recent multisensory experience. *Experimental Brain Research*, *233*, 1983–1991. https://doi.org/10.1007/s00221-015-4259-z

Parise, C. V., Spence, C., & Ernst, M. O. (2012). When correlation implies causation in multisensory integration. *Current Biology*, *22*, 46–49. https://doi.org/10.1016/j.cub.2011.11.039

Park, H., & Kayser, C. (2019). Shared neural underpinnings of multisensory integration and trial-by-trial perceptual recalibration in humans. *eLife*, *8*, e47001. https://doi.org/10.7554/eLife.47001

Passamonti, C., Frissen, I., & Làdavas, E. (2009). Visual recalibration of auditory spatial perception: Two separate neural circuits for perceptual learning. *European Journal of Neuroscience*, *30*, 1141–1150. https://doi.org/10.1111/j.1460-9568.2009.06910.x

Pegado, F., Vankrunkelsven, H., Steyaert, J., Boets, B., & Op de Beeck, H. (2016). Exploring the use of sensorial LTP/LTD-like stimulation to modulate human performance for complex visual stimuli. *PLoS One*, *11*, e0158312. https://doi.org/10.1371/journal.pone.0158312

Phan, M. L., Schendel, K. L., Recanzone, G. H., & Robertson, L. C. (2000). Auditory and visual spatial localization deficits following bilateral parietal lobe lesions in a patient with Balint's syndrome. *Journal of Cognitive Neuroscience*, *12*, 583–600. https://doi.org/10.1162/089892900562354

Radeau, M., & Bertelson, P. (1974). The after-effects of ventriloquism. *Quarterly Journal of Experimental Psychology*, *26*, 63–71. https://doi.org/10.1080/14640747408400388

Ragert, P., Kalisch, T., Bliem, B., Franzkowiak, S., & Dinse, H. R. (2008). Differential effects of tactile high- and low-frequency stimulation on tactile discrimination in human subjects. *BMC Neuroscience*, *9*, 9. https://doi.org/10.1186/1471-2202-9-9

Recanzone, G. H. (1998). Rapidly induced auditory plasticity: The ventriloquism aftereffect. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 869–875. https://doi.org/10.1073/pnas.95.3.869

Sarlat, L., Warusfel, O., & Viaud-Delmon, I. (2006). Ventriloquism aftereffects occur in the rear hemisphere. *Neuroscience Letters*, *404*, 324–329. https://doi.org/10.1016/j.neulet.2006.06.007

Stecker, G. C., Harrington, I. A., & Middlebrooks, J. C. (2005). Location coding by opponent neural populations in the auditory cortex. *PLoS Biology*, *3*, e78. https://doi.org/10.1371/journal.pbio.0030078

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: Current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, *9*, 255–266. https://doi.org/10.1038/nrn2331

Stekelenburg, J. J., Vroomen, J., & de Gelder, B. (2004). Illusory sound shifts induced by the ventriloquist illusion evoke the mismatch negativity. *Neuroscience Letters*, *357*, 163–166. https://doi.org/10.1016/j.neulet.2003.12.085

Strelnikov, K., Rosito, M., & 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

Teder-Sälejärvi, W. A., Di Russo, F., McDonald, J. J., & Hillyard, S. A. (2005). Effects of spatial congruity on audio-visual multimodal integration. *Journal of Cognitive Neuroscience*, *17*, 1396–1409. https://doi.org/10.1162/0898929054985383

Van der Burg, E., Alais, D., & Cass, J. (2015). Audiovisual temporal recalibration occurs independently at two different time scales. *Scientific Reports*, *5*, 14526. https://doi.org/10.1038/srep14526

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., … Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, *25*, 58–76. https://doi.org/10.3758/s13423-017-1323-7

Watson, D. M., Akeroyd, M. A., Roach, N. W., & Webb, B. S. (2019). Distinct mechanisms govern recalibration to audio-visual discrepancies in remote and recent history. *Scientific Reports*, *9*, 8513. https://doi.org/10.1038/s41598-019-44984-9

Woods, T. M., & Recanzone, G. H. (2004). Visually induced plasticity of auditory spatial perception in macaques. *Current Biology*, *14*, 1559–1564. https://doi.org/10.1016/j.cub.2004.08.059

Wozny, D. R., & Shams, L. (2011). Recalibration of auditory space following milliseconds of cross-modal discrepancy. *Journal of Neuroscience*, *31*, 4607–4612. https://doi.org/10.1523/JNEUROSCI.6079-10.2011

Zaidel, A., Turner, A. H., & Angelaki, D. E. (2011). Multisensory calibration is independent of cue reliability. *Journal of Neuroscience*, *31*, 13949–13962. https://doi.org/10.1523/JNEUROSCI.2732-11.2011

Zierul, B., Röder, B., Tempelmann, C., Bruns, P., & Noesselt, T. (2017). The role of auditory cortex in the spatial ventriloquism aftereffect. *NeuroImage*, *162*, 257–268. https://doi.org/10.1016/j.neuroimage.2017.09.002

Zierul, B., Tong, J., Bruns, P., & Röder, B. (2019). Reduced multisensory integration of self-initiated stimuli. *Cognition*, *182*, 349–359. https://doi.org/10.1016/j.cognition.2018.10.019

How to cite this article: Bruns P, Dinse HR, Röder B. Differential effects of the temporal and spatial distribution of audiovisual stimuli on cross-modal spatial recalibration. *Eur J Neurosci*. 2020;00:1–13. https://doi.org/10.1111/ejn.14779