CORTEX 190 (2025) 178–191

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Special Issue 'From Bodies to Spaces': a neurocognitive/neuropsychological...': Research Report

Linking spatial metaphors to body size perception: Different roles of top-down associations and multisensory contributions when mapping auditory cues to finger length

Marte Roel Lesur [a,b,c,*], *Matthew R. Longo* [d] and
Ana Tajadura-Jiménez [a,e,**]

[a] *DEI Interactive Systems Group, Department of Computer Science and Informatics, Universidad Carlos III de Madrid, Spain*
[b] *Association for Independent Research, Switzerland*
[c] *Department of Adult Psychiatry and Psychotherapy, University of Zurich, Switzerland*
[d] *School of Psychological Sciences, Birkbeck University of London, United Kingdom*
[e] *UCL Interaction Centre, University College London, United Kingdom*

ARTICLE INFO

Article history:
Received 29 November 2024
Revised 6 June 2025
Accepted 10 June 2025
Published online 4 July 2025

Keywords:
Multisensory integration
Attention
Body perception
Spatial perception
Sound
Body transformation experiences

ABSTRACT

Temporospatial and semantic multisensory aspects contribute to bodily and spatial perception. An informative paradigm to study this is the Auditory Pinocchio Illusion, in which participants perceive an elongation of their finger upon vertically pulling their finger and hearing a concurrent upward pitch glissando. This arguably relies on anchoring (i.e., associating) the ecologically unrelated upward pitch glissando to the finger and allows to separately assess the role of semantic and multisensory contributions. However, what is needed for this anchoring to occur is unknown. In a first Experiment, we manipulated top-down attention to the finger upon which either an ascending or descending sound would be produced. In a second experiment, we compared how different bottom-up multisensory cues (arising from actions performed on the finger) concurrent to the ascending or descending pitch affected finger length perception. Participants either pulled, touched or stretched their finger. Through a perceptual judgment task of finger landmark localization and questionnaire ratings, we measured participants' perceived finger length in both studies and separately assessed their sensory imagery skills. Our results show that attention alters finger length perception according to questionnaire ratings but not perceptual judgements, while concurrent multisensory signals similarly affect both measures. No relationship between these effects and participants' sensory imagery was found. We suggest that while top-down associations between pitch and verticality are necessary

---

This article is part of a special issue entitled: From Bodies to Spaces published in Cortex.

[*] *Corresponding author*. DEI Interactive Systems Group, Department of Computer Science and Informatics, Universidad Carlos III de Madrid, Spain.

[**] *Corresponding author*. DEI Interactive Systems Group, Department of Computer Science and Informatics, Universidad Carlos III de Madrid, Spain.

E-mail addresses: mroel@inf.uc3m.es (M. Roel Lesur), atajadur@inf.uc3m.es (A. Tajadura-Jiménez).

https://doi.org/10.1016/j.cortex.2025.06.014

and affect questionnaire ratings, they are not sufficient to affect perceptual judgements. Bottom-up somatosensory cues seem to be additionally needed to impact such judgements in this illusion.

1. Introduction

In recent years, many illusions illustrating and attempting to understand the malleability of one's own body perception have been documented (Lackner, 1988; Ehrsson, 2012; Kilteni, Maselli, Kording, & Slater, 2015; Tajadura-Jiménez, Väljamäe, et al., 2017; Roel Lesur et al., 2020a; Tajadura-Jiménez, Fairhurst, & Deroy, 2022). These ingenious procedures leading to body transformation experiences have offered insights into fundamental scientific questions as well as potential therapies and other applications (Dijkerman & Lenggenhager, 2018; Kilteni et al., 2015). Alterations in one's body perception in turn impact cognition at large (Bolt et al., 2021; Dijkerman & Lenggenhager, 2018; Maister, Slater, Sanchez-Vives, & Tsakiris, 2015), including spatial aspects such as one's body size (Hoort, Guterstam, & Ehrsson, 2011; Tajadura-Jiménez et al., 2012, 2018) as well as distance and size perception of extrapersonal objects or people (Hoort et al., 2011; Lenggenhager, Mouthon, & Blanke, 2009; Roel Lesur, Bolt, Saetta, & Lenggenhager, 2021; Tajadura-Jiménez, Banakou, et al., 2017). Some of these perceptual changes are produced under peculiar circumstances that might give insight into the functioning of bodily and spatial cognition.

In the Auditory Pinocchio illusion, participants perceive their finger length increasing when an upward pitch glissando is heard concurrent to pulling their vertically positioned index finger. The illusion emerges when the finger is vertically extended while hidden from view, and either the experimenter (Nava & Tajadura-Jiménez, 2020) or participants themselves (Tajadura-Jiménez, Vakali, et al., 2017) pull it upwards while an upward pitch glissando is concurrently produced (on headphones, with the aid of a sensor which detects when touch is applied to the finger in order to trigger the playback of the sound). A change in perceived finger length has been measured both through questionnaires and behavioural measures (Nava & Tajadura-Jiménez, 2020; Tajadura-Jiménez, Vakali, et al., 2017). For the latter, across several repetitions, participants either passively (i.e., without moving) select the location of their fingertip and finger base by saying stop when the experimenter passes a ruler by the respective location (Tajadura-Jiménez et al., 2017b), or participants themselves actively point at each of these locations (Nava & Tajadura-Jiménez, 2020).

The illusion builds on a prior (and potentially universal) spatial association between pitch and height or size (for a review see Deroy, Fernandez-Prieto, Navarra, & Spence, 2018; see also Eitan, Schupak, Gotler, & Marks, 2014; Evans & Treisman, 2010; Newbold, Gold, & Bianchi-Berthouze, 2020; Spence, 2019; Parise, Knorre, & Ernst, 2014). This crossmodal association is arguably based on linking the perceived direction of the glissando and the vertical pulling to create a sense of stretching the finger. The paradigm is based on conceptual mappings articulated in the vast literature on cross-modal correspondences (Spence, 2011, 2019; Walker & Walker, 2016). These mappings refer to the linking of different perceptual features for which the same conceptual label exists (e.g., the concepts *low* and *high* and their respective relation to pitch and visuospatial elevation). The subjective and recurrent pairing or association between an arbitrary sound and the finger has been referred to as anchoring in previous work on the Auditory Pinocchio illusion (see Tajadura-Jiménez, Vakali et al., 2017). This illusion relies on an auditory cue ecologically unrelated to the body (i.e., an arbitrary pitch shift that is not naturally occurring) and thus offers unique possibilities to understand how arbitrary cues might be anchored to one's body and affect bodily and spatial perception.

However, the conditions that enable this anchoring of sound to the body remain unclear. Research has shown that the Auditory Pinocchio illusion arises when a finger is pulled distally with the tip facing upwards or downwards in concordance with the ascending pitch shift, but not with a constant tone nor a descending pitch change. The alteration occurs only towards a lengthened but not a shortened finger, even when it is pushed proximally and together with a descending pitch (Nava & Tajadura-Jiménez, 2020; Tajadura-Jiménez, Vakali, et al., 2017). Such unidirectional shift towards an enlarged body part is consistent with existing literature on alterations of body perception beyond sound (Giurgola, Pisoni, Maravita, Vallar, & Bolognini, 2019; de Vignemont, Ehrsson, & Haggard, 2005; Marino, Stucchi, Nava, Haggard, & Maravita, 2010; Pavani & Zampini, 2007; Lackner, 1988). For example, Marino et al. (2010), found changes in grasping behaviour when increasing the image of participants' hand but not when shrinking it, arguing that enlargement (e.g., during ontogenetic growth) is ecologically more common than shrinking. Furthermore, the illusion seems to be limited to verticality, as it does not occur when the target finger is positioned horizontally and pulled distally in the same axis (Nava & Tajadura-Jiménez, 2020), further showing the vertical specificity of this effect.

While these contingencies give insights into the scope and limits for the Auditory Pinocchio illusion to emerge, the relative role of top-down priors (e.g., the link between pitch and verticality) and bottom-up sensory signals is unclear. In the literature on bodily illusions, linking information from one modality to the other is generally attributed to three sources. First, to the temporal synchrony between spatially incongruent multisensory cues which is usually thought of as a bottom-up factor (Botvinick & Cohen, 1998; Ehrsson, 2012; Roel Lesur, Gaebler, Bertrand, & Lenggenhager, 2018; Roel Lesur, Weijs, et al., 2020; Vroomen & De Gelder, 2004; Bertelson & Aschersleben, 1998; Caclin, Soto-Faraco, Kingstone, & Spence, 2002). A well-known example is the rubber hand illusion, in which seeing a rubber hand being stroked simultaneously to one's own occluded hand facing in the same direction results in the feeling that the rubber hand is one's own (Botvinick & Cohen, 1998; Riemer, Trojan, Beauchamp, & Fuchs, 2019). Arguably, the felt touch is visually captured at the rubber hand's location resulting in this experience. Second, a degree of spatial congruency (Kammers, Longo, Tsakiris, Dijkerman, & Haggard, 2009; Tsakiris & Haggard, 2005; Rohde, Luca, & Ernst, 2011). For example, if a different finger is stroked on the rubber hand and participants' own hand there is no illusion despite temporal synchrony (Kammers et al., 2009); also, a 90° rotated rubber hand also does not produce the illusion when the same finger is stroked (Tsakiris & Haggard, 2006). Spatial congruence (e.g., the directionality of touch, or correspondence between the seen and felt finger) is generally considered a bottom-up factor. And lastly, semantic congruency: If instead of a rubber hand, participants see an unrelated object such as a stick (Tsakiris & Haggard, 2005), the illusion does not work, an aspect clearly reliant on prior knowledge and expectations and thus considered top-down (Apps & Tsakiris, 2014). It should be noted that despite the apparently clear distinction of these aspects, they are often interrelated. In the Auditory Pinocchio illusion, the semantic congruency between the pitch change and vertical extension seems important, but whether this drives the illusion or if indeed multisensory signals are necessary, and if so, which type of signals are required, is unclear.

To assess this, we ran two experiments employing a perceptual judgment task and questionnaire ratings. Experiment 1 evaluated whether attending a vertically positioned finger and hearing a subsequent sound would be enough for anchoring the sound and producing the illusion in the absence of any direct sensory stimulation to the finger. Since most literature on bodily illusions supports the role of bottom-up temporal synchrony for linking plausible but conflicting cues across sensory modalities (Apps & Tsakiris, 2014; Blanke, 2012; De Vignemont, 2017; Ehrsson, 2012), we expected no effect of attention (i.e., attended versus unattended finger) for the illusion to emerge. However, given the peculiarity of the Auditory Pinocchio illusion (i.e., anchoring an arbitrary non-naturalistic sound to one's finger), it could be that it is primarily driven by a tendency to match experiences to expectations (e.g., potentially linked to cultural associations as the use of auditory glissandos in cartoons) also referred to as phenomenological control. Phenomenological control, which is linked to mental imagery, has been argued to contribute to related bodily illusions (Dienes et al., 2022; Lush et al., 2019, 2020). While this argument has been questioned (Ehrsson, Fotopoulou, Radziun, Longo, & Tsakiris, 2022), given its specificity, the Auditory Pinocchio Illusion could be plausibly driven by such factors. To discard this, we studied the role of top-down attentional factors in the illusion. To further assess whether one's prior capacity to imagine across sensory modalities might be involved in anchoring the sound to the finger (potentially via suggestibility), we measured participants' sensory imagery skills through a questionnaire (Andrade, May, Deeprose, Baugh, & Ganis, 2014; Pérez-Fabello & Campos, 2020). Sensory imagery refers to one's capacity to portrait and recreate mental images across sensory modalities with different degrees of vividness (Andrade et al., 2014; Pérez-Fabello & Campos, 2020) and has been linked to suggestibility (Palmer & Field, 1968).

In Experiment 2, to confirm whether any multisensory cues related to the finger are enough to anchor the sound to one's finger and which coupling might yield a stronger illusion of finger elongation, we compared how different multisensory combinations contribute to the illusion. In previous studies, participants' fingers were pulled either by themselves or the experimenter, thus involving touch, stretching and pulling. Whether each of these cues might differentially affect the anchoring of the sound, and in turn the illusion, is unknown. Participants either pulled the target (left index) finger with their right index and thumb (*pulling* condition), or grabbed it without pulling (*touching* condition), or stretched the finger upward without any additional touch (i.e., extending the muscles of their own left index finger; *stretching* condition). Each of these actions activated a sensor to trigger the sound. We expected the greatest effect in the pulling condition due to the greatest signal redundancy, followed by the stretching one due to the vertical direction of the action and the above-mentioned association of pitch and verticality (see Deroy et al., 2018). And lastly a smaller effect for the touching condition given that the auditory cue would be the only sensory signal indicating vertical displacement.

Given the separate assessment of top-down and bottom-up signals in our respective experiments, we explored whether correspondingly differing results for perceptual judgements and questionnaire ratings might be found.

2. Methods

2.1. Participants

Using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007), we performed a power analysis based on Experiments 1 and 2 from Tajadura-Jiménez, Vakali et al. (2017), measuring the Auditory Pinocchio illusion (Cohen's $d = .591$, and .583, respectively, for an average of .587). The results called for a sample of 25 participants for a power of .8, or 33 participants for a power of .9. Given the novel manipulations intended for this study, the additional levels within each factor, the different number of trials, and the lack of previous data, we recruited 40 participants from the Universidad Carlos III de Madrid. All participants took part in both experiments. We recruited on the basis of being between 18 and 60 years of age, speaking Spanish, with no previous history of psychiatric or neurological disorders, and not having difficulties with distinguishing between left and right. Ethical approval was granted by the University (approval number 101002711), participants gave their informed consent, and a €10 compensation was offered for taking part in both experiments. Data from two participants was discarded from the whole experiment; one due to data loss, and another because the participant reported falling asleep during the task. Data from an additional participant was discarded from Experiment 2 due to malfunctioning of the pressure sensor. Lastly, EMG sensor data from two additional participants from Experiment 1 were further discarded due to logging issues. Resulting in $N = 38$ for Experiment 1 (27 female, 2 non-binary; age: $M = 26.2$, $SD = 9.2$) and $N = 37$ for Experiment 2 (27 female, 2 non-binary; age: $M = 25.8$, $SD = 9$), and $N = 36$ for the EMG sensor data analyses of Experiment 1.

2.2. Procedure

2.2.1. Design Experiment 1

In block 1 (perceptual judgement task), 40 trials were presented in a random order, 10 for each combination of sound (ascending or descending) and attention (attended or unattended finger). Each side was attended and unattended an equal number of times. The order of the trials was pseudorandomized. Upon completion, block 2 followed. In it, participants filled-in a finger elongation and embodiment questionnaire with the order of conditions semi-counterbalanced (four different orders, each with a different combination of starting side, attention, and sound were presented across participants). After 10% of the trials a certainty question was presented at a random order within the task to ensure that participants were attending the correct side.

2.2.2. Design Experiment 2

For block 1 (perceptual judgement task) a total of 12 repetitions were presented per condition (touching, pulling, or stretching), with half of the trials for each of the sounds (descending or ascending). The order was randomized. Afterwards, in block 2, for each combination of action (touching, pulling, stretching) and sound (ascending, descending) a finger elongation and embodiment questionnaire was administered so that participants completed the questionnaire six times in a random order.

2.2.3. Preparation

See Fig. 1 for an overview of the experimental design and procedure. Before attending the experimental session, participants filled in an online questionnaire (see section 2.4.3 Sensory imagery). To confirm their capacity to distinguish between left and right, they were asked to lift either their left or right arm as soon as they heard the experimenter instructing either side and bring the arm down afterward. A total of 10 trials for each side were presented in a random order for this side detection task. Only participants with less than 2 errors were considered for the remainder of the study. In our sample, participants missed no trials except for one missing a single early trial but adjusting for the remainder of the task.

Participants' index fingers were measured from the tip to the proximal digital crease (the line connecting the palm and the finger). They were asked to look away during the measurement. After this, their proximal digital creases were marked in red and shown as a visual reference for what they should attempt to locate (without seeing) during the experiment. For the tip of the finger, they were instructed to focus on the farthest position of their fingertip. Their hands were placed resting on the corresponding location in front of a computer for the procedure of Experiment 2 (see Fig. 2, right panel). Instructions for the experiment followed.

Participants completed Experiment 1 before Experiment 2; however, the latter was explained first to ensure the functionality of all the sensors. Detailed guides on each of the actions necessary for the experiment were presented,

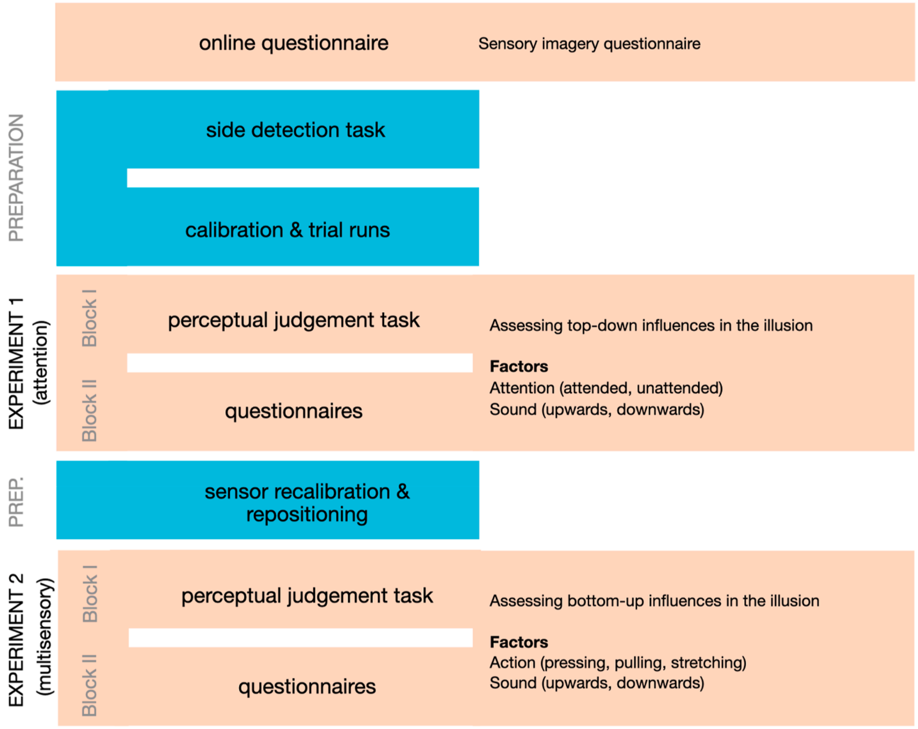

Fig. 1 – Experimental design and procedure. The order of the sections is presented from top to bottom with the data gatherings sections in green and preparation sections in grey. All participants took part in both experiments.

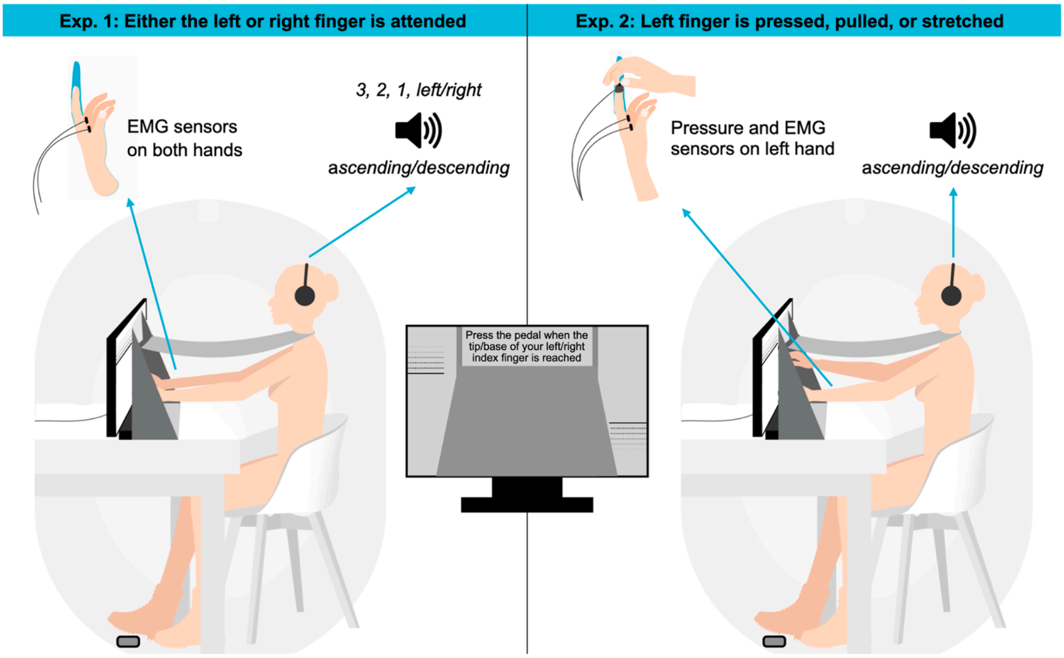

Fig. 2 – Experimental setup for Experiment 1 (left panel) and 2 (right panel). The computer depicted in the middle illustrates the finger length estimation task for Experiment 1 (i.e., for each hand). In Experiment 2 the finger length estimation task was only performed for the left finger (as in Tajadura-Jiménez, Vakali et al., 2017).

correspondingly for stretching, touching, and pulling. For the experimenter to assess whether their actions would activate the sensors as expected, a button would light up on the computer upon each action. Once this was confirmed, participants' hands were occluded with a hairdressing apron attached to the top of the platform on one side and to their neck on the other side (Fig. 2, right panel). Participants were then instructed to activate a sound (a 350 Hz sine wave) with each action four times; if any issues were found with the sensors they were recalibrated. This was for participants to acquaint themselves with the task as well as to anchor the sound to the finger, as done in previous studies (Nava & Tajadura-Jiménez, 2020; Tajadura-Jiménez, Vakali, et al., 2017).

Participants were then prompted to place both hands in the corresponding location for Experiment 1 (Fig. 2, left panel), the pressure sensor was removed, their hands hidden, and the instructions were presented. Participants heard on the headphones a recorded countdown stating *three, two, one,* and either *left* or *right*, upon which they had to attend to their index finger on the indicated side. A pitch would then sound 500 ms after the side was instructed. Following the pitch, an instruction appeared on the screen indicating to prepare for estimating either the right or left finger length. The experimenter then triggered a trial of the Finger length estimation task (see section 2.4.1) for which a line appeared on the corresponding top side of the screen and moved until it reached participants' fingers' tips and bases (Fig. 2, middle) which indicated that they must press a pedal. For each trial, participants first heard the instruction indicating the index finger (left or right) to attend while the sound (ascending or descending pitch) was presented and then saw on the screen the index finger (left or right) whose size had to be estimated. Participants completed this task twice for each combination of side and measured finger so that in half of the trials the attended finger did not correspond to the measured finger (corresponding to an *unattended* finger trial). They were informed to be attentive to the instructed side as in 10% of the times the software asked them which side was attended to ensure they were not responding at random. Upon completing all the trials, a second block started. For each of the combinations of attention (attended and unattended) and sound (ascending and descending) a questionnaire was presented. Upon completion, participants could take a 5 min break before repositioning the sensors and arm placements for Experiment 2. Their hands were readjusted, a cushion was placed under their right elbow for convenience, and the correct functioning of the electrodes was ensured for each of the actions involved in Experiment 2 (*touching*, *pulling* or *stretching*).

2.3. *Apparatus*

The two auditory stimuli used were 2000 ms long sine waves with the same amplitude which was played back at the same level to all participants ensuring this through the computer level meter. A linear ramp of the same duration produced a glissando. For the ascending sound the pitch began at 700 Hz and ended at 1200 Hz while for the descending one it began at 1200 Hz and ended at 700 Hz. Acoustic stimulation was presented to the participants via headphones (Sennheiser HD 2.30G, Sennheiser, Germany). A 350 Hz sine wave was used as a neutral pitch during calibration and served as an initial anchoring of the pitch to the finger (Tajadura-Jiménez, Vakali, et al., 2017; c.f., Walker & Walker, 2016). The countdown sounds announcing the beginning of the trial for Experiment 1 (indicating *three, two, one,* and either *left* or *right*) had the same duration and length for each word. Stimuli presentation, randomization and data recording was performed using Unity 2021 running on a Windows computer (Windows 10, 16 GB RAM, AMD Ryzen 5 2600 6-core processor at 3400 Mhz). The display on which participants laid their hands showed a fixation mark to be looked at during the trials. This was possible via a customized apron that covered the view over the arms but would leave visible a portion of the screen (see Fig. 2).

A pressure sensor (force, FSR sensor, Biosignalsplux, Portugal), a pedal (Footswitch, Biosignalsplux, Portugal) to respond during the finger length estimation task, and an electromyography (EMG, Bitalino, Portugal) sensor were connected to the computer via Bluetooth utilizing a Bitalino Core (Biosignalsplux, Portugal). In Experiment 1, the EMG was gathered to ensure that participants were not extending one index finger more than the other (Tajadura-Jiménez, Vakali, et al., 2017). In Experiment 2, the pressure sensors tracked the pressure on the target finger for activating sounds (in the *touching* condition) while the EMG used the muscular extension for the same purpose. For the EMG, two 2.4 mm electrodes were placed on the dorsal part of each hand, one about 1 cm below the knuckle and another one about 2 cm above the wrist on the same axis of the index finger. Another electrode was placed near the elbow, on the bony part of the lower arm. For the pressure sensor, Velcro straps were used to comfortably fit the sensor on the distal phalanx of the left index finger. An additional circuit was set to ensure that participants did not move their fingers during Experiment 1. The device consisted of an Arduino Mini and infrared sensors (which detect distance by emitting an infrared signal and reading the reflected energy to measure distance) to give a .5 cm range beyond which an alarm sounded, prompting the experimenter to adjust participants' fingers. Visual stimuli were presented on an AOC Q32V4 monitor (2560 × 1440 pixels; 70 × 39.5 cm).

2.4. *Measures*

2.4.1. *Perceptual judgement task*

Participants placed the measured finger lying vertically on a plastic panel located on a large computer screen. This was occluded from their view using an apron. On the top corner corresponding to the measured finger's side a line appeared and moved downwards at a speed of 2.63 cm per sec. Participants had to press a pedal once they felt the line reached their fingertip. The line then appeared on the bottom of the screen and moved upwards at the same speed. Participants had to press the pedal when it seemed to reach their finger base. This digital version is an improvement to the method used by Nava and Tajadura-Jiménez (2020) and Tajadura-Jiménez, Vakali, et al. (2017; see also Mancini, Longo, Kammers, & Haggard, 2011) to assess perceptual judgments of finger landmark locations in the Auditory Pinocchio illusion. It reduces the measurement error of the experimental observer and removes error in the mechanical movement of the clip by the experimenter (instead of an automated system). It further has a higher resolution of .027 cm (the size of each pixel) compared to the previous .5 cm resolution.

2.4.2. *Finger elongation and embodiment questionnaire*

An adaptation of the questionnaire used by Tajadura-Jiménez, Vakali, et al., 2017 was used to assess the phenomenal quality of the experience (see Tables 1 and 2). The questionnaires were presented in Spanish. Two new items regarding embodiment of the sound were included, one of them is: "it felt as if the sound was produced by my own finger" and is meant to explore whether the sound was perceived as part of the body rather than as an external cue interacting with bodily experience. The other item was for Experiment 1 "it felt as if attending my finger produced the sound" and for Experiment 2 "it felt as if the gesture produced the sound" and was meant to assess whether in fact participants linked either of these actions to producing the sound. Though the term "gesture" was used in the Spanish questionnaires, it corresponds to the actions for Experiment 2.

An additional item also used by Tajadura-Jiménez, Vakali, et al. (2017; see also de Vignemont et al., 2005) was composed of nine images each depicting a hand with an index finger of different elongations from which participants had to select the one better representing their perception of their own hand while hearing the sound.

Table 1 – Results of ART ANOVAs for the finger elongation and embodiment questionnaire of Experiment 1.

Item	Question	Attention			Sound			Interaction		
		F (df)	p	η_p^2	F (df)	p	η_p^2	F (df)	p	η_p^2
1	*… attending my finger produced the sound*	11.29 (1,37)	**.014**	.234	.27 (1,37)	.821	.007	.42 (1,37)	.731	.011
2	*… my finger was longer*	7.83 (1,37)	**.034**	.175	10.27 (1,37)	**.018**	.217	3.82 (1,37)	.143	.094
3	*… my finger was shorter*	.71 (1,37)	.628	.019	7.86 (1,37)	**.034**	.175	14.51 (1,37)	**.008**	.282
4	*… my finger was elevating*	8.59 (1,37)	**.032**	.188	6.51 (1,37)	**.045**	.15	6.5 (1,37)	**.045**	.149
5	*… my finger was descending*	1.09 (1,37)	.509	.029	22.12 (1,37)	**.008**	.374	15.42 (1,37)	**.008**	.294
6	*… my finger was elongating*	7.65 (1,37)	**.034**	.171	20 (1,37)	**.008**	.351	6.86 (1,37)	**.045**	.156
7	*… my finger was shrinking*	3.11 (1,37)	.201	.078	6.22 (1,37)	**.045**	.144	6.25 (1,37)	**.045**	.145
8	*… I could not tell how long my finger was*	2.41 (1,37)	.258	.061	.15 (1,37)	.847	.004	2.03 (1,37)	.309	.052
9	*… I could not locate my finger base*	.71 (1,37)	.628	.019	.05 (1,37)	.888	.001	.14 (1,37)	.847	.004
10	*… I could not locate my fingertip*	.15 (1,37)	.847	.004	.6 (1,37)	.645	.016	1.6 (1,37)	.389	.042
11	*… my finger sensations were unexpected*	.03 (1,37)	.888	.001	.06 (1,37)	.888	.002	.2 (1,37)	.847	.005
12	*… my finger was not mine*	0 (1,37)	.981	<.001	2.87 (1,37)	.208	.072	.04 (1,37)	.888	.001
13	*… my finger was as if numb*	.08 (1,37)	.888	.002	.04 (1,37)	.888	.001	1.48 (1,37)	.404	.039
14	*… the sound was produced by my own finger*	13.52 (1,37)	**.008**	.268	2.99 (1,37)	.203	.075	.64 (1,37)	.641	.017

FDR corrected P-values are shown. Significant findings are presented in bold. The questions always began with, "*while I listened to the sound I felt that …*"

Table 2 – ART ANOVAs for the finger elongation and embodiment questionnaire of Experiment 2.

Item	Question	Action			Sound			Interaction		
		F (df)	p	η_p^2	F (df)	p	η_p^2	F (df)	p	η_p^2
1	*… the gesture produced the sound*	.93 (2,72)	.576	.025	.12 (1,36)	.827	.003	.99 (2,72)	.576	.027
2	*… my finger was longer*	22.61 (2,72)	**.004**	.386	16.88 (1,36)	**.004**	.319	2.31 (2,72)	.223	.06
3	*… my finger was shorter*	11.71 (2,72)	**.004**	.245	14.24 (1,36)	**.004**	.283	.32 (2,72)	.827	.009
4	*… my finger was elevating*	1.7 (2,72)	**.004**	.229	18.33 (1,36)	**.004**	.337	3.75 (2,72)	.078	.094
5	*… my finger was descending*	3.34 (2,72)	.106	.085	17.27 (1,36)	**.004**	.324	.49 (2,72)	.777	.014
6	*… my finger was elongating*	37.39 (2,72)	**.004**	.509	21.06 (1,36)	**.004**	.369	.24 (2,72)	.827	.007
7	*… my finger was shrinking*	9.77 (2,72)	**.004**	.213	1.94 (1,36)	**.008**	.233	.51 (2,72)	.777	.014
8	*… I could not tell how long my finger was*	1.65 (2,72)	.382	.044	.76 (1,36)	.576	.021	1.29 (2,72)	.472	.035
9	*… I could not locate my finger base*	.33 (2,72)	.827	.009	1 (1,36)	.522	.027	1.59 (2,72)	.385	.042
10	*… I could not locate my fingertip*	2.13 (2,72)	.254	.056	3.32 (1,36)	.170	.084	.25 (2,72)	.827	.007
11	*… my finger sensations were unexpected*	.47 (2,72)	.777	.013	8.41 (1,36)	**.019**	.189	.15 (2,72)	.878	.004
12	*… my finger was not mine*	.25 (2,72)	.827	.007	1.38 (1,36)	.434	.037	.07 (2,72)	.934	.002
13	*… my finger felt as if numb*	2.72 (2,72)	.168	.07	.24 (1,36)	.777	.007	6.14 (2,72)	**.011**	.146
14	*… the sound was produced by my own finger*	.85 (2,72)	.603	.023	6.6 (1,36)	**.045**	.155	3.28 (2,72)	.106	.083

FDR corrected P-values are shown. Significant P-values are presented in bold. The questions always began with, "*while I listened to the sound I felt that …*"

2.4.3. *Sensory imagery*

Participants filled in the Spanish version of the Plymouth sensory imagery questionnaire (Pérez-Fabello & Campos, 2020) consisting of 35 questions assessing sensory imagery across sensory modalities. The questionnaire includes items addressing auditory, visual, tactile, olfactory, gustatory and motor imagery.

2.4.4. *Sensor data*

To assess whether participants indeed did not move or press the relevant finger differently between conditions in each of the experiments, we measured EMG and pressure sensor data for each trial and relevant finger. Sensor data was gathered at a 100 Hz sampling rate for the duration of the triggered sound.

2.5. *Data processing and statistical analysis*

Data was collected in Unity and output into CSV files to later be processed using R (version 4.2.2). Descriptive statistics, plots, and aligned ranks transformed analyses of variance (ART-ANOVA; Kay & Wobbrock, 2020; Wobbrock, Findlater, Gergle, & Higgins, 2011) for nonparametric factorial analyses (for questionnaire data) were conducted using R. Shapiro–Wilk normality tests and visual inspection were performed. For the questionnaires, P-values were adjusted for multiple comparisons according to the Benjamini-Hochberg or False Discovery Rate (FDR) procedure (Abdi, 2007). To analyse data from the sensors and finger length estimation tasks Bayesian Repeated Measures (RM) ANOVAs were conducted using JASP (version .18.1), the equivalent frequentist analyses are reported in the Supplementary Materials. Bayesian analyses quantified the relative predictive performance of two rival hypotheses, allowing to quantify the performance of the null hypothesis (Van Den Bergh et al., 2020; van Doorn et al., 2021). A Bayes Factor (BF_{10}) < 1/3 is indicative of moderate evidence in favour of the null hypothesis (H_0) while a Bayes Factor >3 is considered moderate evidence to support the alternative hypothesis (H_1). For Bayesian RM ANOVAs the different models were compared to the null model including random intercepts and slopes.

Bayesian correlation analyses were performed with JASP (version .18.1). We assessed whether Plymouth sensory imagery questionnaire scores were correlated with the difference in participants' estimated finger length between conditions for each of the experiments and the sum of the two open questions regarding finger length estimation (items 2 and 3, in Tables 1 and 2). For these tests, the auditory, bodily and overall scores from the Plymouth sensory imagery scores were considered. These categories were chosen given the auditory and bodily nature of the tasks, together with a more general score for mental imagery.

For the finger length estimation task, only data within 3 standard deviations (SD) from the general mean for both the tip and the base were considered, and otherwise substituted by the participants' mean for that condition. This approach for rejecting outliers is the same used in previous research on the Auditory Pinocchio Illusion (Navas-León et al., 2023; Tajadura Tajadura-Jiménez, Vakali, et al., 2017) and resulted in 1.18% of trials for the Experiment 1 and 3.77% of trials for Experiment 2. The same criterion of 3 SD from the mean was applied for the sensor data when calculating the mean for each participant and condition (see Supplementary Materials).

3. Results

3.1. *Results Experiment 1*

3.1.1. *Perceptual judgement task*

As hypothesized, a Bayesian RM ANOVA showed moderate evidence that there was no effect of sound (i.e., ascending or descending pitch; $BF_{10} = .21$) and strong evidence for a lack of interaction between sound and attention ($BF_{10} = .06$). This suggests that sound alone is not enough to elicit the perceptual change according to this measure in the Auditory Pinocchio illusion even when combined with a top-down association to the finger (i.e., attention). As for the main effect of attention, the results did not reach our threshold of moderate evidence in favour or against the null hypothesis (BF_{10} = 1.157) thus yielding inconclusive results. More data might be necessary to confirm whether not attending the finger (mean = 8.40 cm, SD = 2.26 cm) leads to an increase in perceived finger size compared to attending it (mean = 8.14 cm, SD = 2.19 cm), however this question is outside the aim of our study.

Additional analyses explored the degree of correlation with the questionnaire assessments despite the lack of effect, correspondingly yielding inconclusive Bayesian evidence (see Supplementary Materials). Sensor data showed no interaction between sound and condition, suggesting no relevant differences in finger extension for the relevant comparisons (see Supplementary Materials).

3.1.2. *Finger elongation and embodiment questionnaire*

Results from the ART ANOVAs for each of the questionnaire items are presented in Table 1. Descriptive statistics can be found in the Supplementary Materials. To summarize the questionnaire results, there was an effect of attention for questions referring to feeling that the sound was produced by the finger (item 1), that the finger was longer (item 2; Fig. 3b), that it was elevating (item 4), elongating (item 6), and that it was produced by participants' own finger (item 14). An effect of sound was found for the feeling that the finger was longer (item 2; Fig. 3b), shorter (item 3, Fig. 3b), elevating (item 4), descending (item 5), elongating (item 6), and shrinking (item 7). These findings show that participants reported feeling a spatial change in the corresponding sound direction even despite the involvement of additional multisensory cues. Finally, an interaction between attention and sound was found for the feeling that the finger was shorter (item 3; Fig. 3b), elevating (item 4), descending (item 5), elongating (item 6), and shrinking (item 7), suggesting that for some items the joint involvement of attention and auditory cues has an effect.

At last, for the item in which participants selected among 9 images each depicting a different finger size on a Likert scale, there was no effect of attention ($F(1, 37) = .09$, $p = .768$, $\eta_p^2 = .002$), but an effect of sound ($F(1, 37) = 6.87$, $p = .013$, $\eta_p^2 = .157$) and an interaction of sound and attention ($F(1, 37) = 7.5$, $p = .009$, $\eta_p^2 = .168$) (see Fig. 3c).

Figure 3b and c respectively show the mean of items 2 and 3 (inverted) of the questionnaire, and the item in which participants had to select the image best representing their finger. This illustrates that participants did report a change in finger size corresponding to the direction of the sound when attending the finger. The main effect of attention for items 1 and 14 (see Fig. 4) suggests that participants did attribute the sound to the finger more in the attended than in the unattended conditions.

3.1.3. *Sensory imagery*

Given that no differences were found between sound conditions for the perceptual judgement task no further correlation with sensory imagery was performed.

For the questionnaires, we calculated the mean responses for the items respectively referring to perceiving the finger as longer and shorter (7-response) and subtracted the mean response for all descending conditions from the mean response for all ascending conditions (i.e., attended and unattended). Bayesian Kendall's Tau-b correlations were performed between the difference in self-reported size and three dimensions of the Plymouth's sensory imagery questionnaire. No correlation was found for either auditory ($\tau_b = .004$, $BF_{10} = .218$), bodily ($\tau_b = .135$, $BF_{10} = .411$), and general imagery ($\tau_b = .075$, $BF_{10} = .366$).

3.2. *Results experiment 2*

3.2.1. *Perceptual judgement task*

We hypothesized that actions involving different multisensory combinations would yield differences for this judgement. However, according to Bayesian RM ANOVAs, we found moderate evidence for a main effect of sound ($BF_{10} = 4.212$) but not of action ($BF_{10} = .101$) nor the interaction of sound and

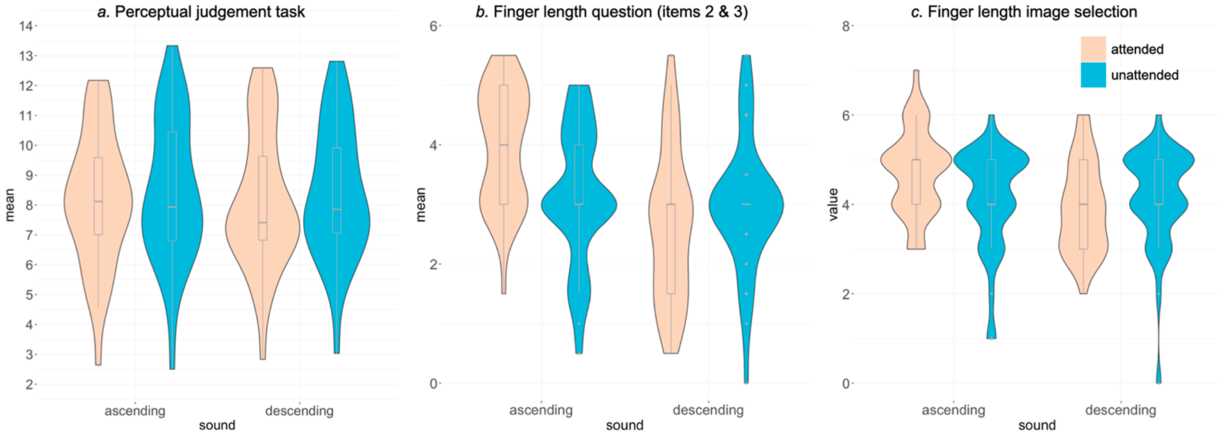

Fig. 3 – Plots showing means, standard errors, central tendencies, and distribution respectively for: (*a*) the mean estimated finger length, (*b*) the mean of questions 2 and 3 of the finger elongation and embodiment questionnaire, and (*c*) the image best representing participants felt finger size pertaining to the same questionnaire of Experiment 1. The evaluated factors are attention (attended/unattended) and sound (ascending/descending).

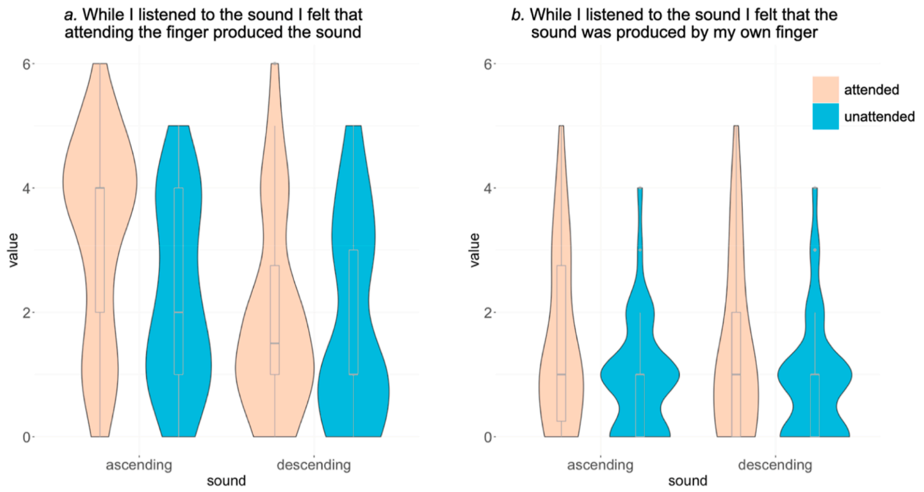

Fig. 4 – Plots showing means, standard errors, central tendencies, and distribution respectively for: (*a*) item 1 referring to the feeling that attending the finger produced the sound, and (*b*) 14, corresponding to the feeling that sound was produced by participants' finger. For both, a main effect of attention was found. The evaluated factors are attention (attended/ unattended) and sound (ascending/descending).

action (BF_{10} = .055). The effect of sound with concurrent stimulation of the finger on perceived finger length confirms previous research (Nava & Tajadura-Jiménez, 2020; Tajadura-Jiménez, Vakali, et al., 2017), however it seems that neither the type of sensory signals involved in addition to sound, nor the degree of signal redundancy play a different role according to this measure. The difference between ascending and descending sounds for each of the actions did reproduce the illusion according this task.

Results from the sensor data (Supplementary Materials) provide moderate evidence that there were no effects of action and sound nor their interaction for finger extension. For finger pressure, we found no clear evidence for a relationship with any of our independent factors. The reason for this is unclear and beyond the scope of our study.

3.2.2. *Finger elongation and embodiment questionnaire*

Statistical results from the ART ANOVAs for each of the questionnaire items are presented in Table 2 below. Descriptive statistics can be found in the Supplementary Materials. There was an effect of action for items referring to feeling that the finger was longer (item 2), shorter (item 3), elevating (item 4), elevating (item 5), descending (item 6), elongating (item 7), and shrinking (item 8). An effect of sound was found for the feeling that the finger was longer (item 2), shorter (item 3), elevating (item 4), descending (item 5), elongating (item 6), shrinking (item 7), for unexpected sensations (item 12), and for the feeling that the sound was produced by the gesture (hereby referred to as action, item 14). As for the interaction between action and sound, an effect was found only for the items referring to the finger feeling numb (item 13) and the sound being produced by the finger (item 14). For the last item (not on the table) involving 9 images depicting a different finger sizes on a Likert scale, there was an effect of action ($F(1, 72) = 17.33$, $p < .001$, $\eta_p^2 = .325$), sound ($F(1, 36) = 7.43$, $p = .01$, $\eta_p^2 = .171$) and no interaction of sound and action ($F(1, 37) = 2.02$, $p = .014$, $\eta_p^2 = .053$).

Figure 5b and c shows a clear effect of sound for each of the actions, where participants felt their finger as longer during the ascending condition. Interestingly, participants generally underestimated their finger size in the touching condition compared to stretching or pulling. We further plot in Fig. 6a responses for the item referring to unexpected sensations (item 12) showing that participants generally felt stranger sensations for the descending sound. We further computed Bonferroni corrected post hoc Wilcoxon signed rank tests for item 12 (Fig. 6b) referring to the feeling of numbness. The results reveal that for ascending sounds there is a significantly greater degree of numbness sensations for the touching condition compared to the pulling ($W = 111$, $p = .006$, $r = -.48$) and the stretching ($W = 207.5$, $p = .002$, $r = -.54$). For the descending sound, no significant differences were found between touching and pulling ($W = 73$, $p = 1$, $r = -.04$) or stretching ($W = 104$, $p < .921$, $r = -.12$).

3.2.3. *Relation between perceptual judgments and questionnaire ratings*

We assessed whether the two measures were correlated by subtracting the mean of all descending conditions (i.e., for stretching, pulling and touching) from the mean of all ascending conditions for both the perceptual judgement task and the questionnaire ratings. For the questionnaire, the mean of the items referring to perceiving the finger as longer and shorter (items 2 and 3 respectively; for the latter the inverted value) was considered. A Bayesian Kendall's Tau-b correlation analysis showed strong evidence of correlation between these two measures ($\tau_b = .396$, $BF_{10} = 35.711$).

3.2.4. *Sensory imagery*

As in Experiment 1, to assess whether sensory imagery would be correlated to the differences in the perceptual judgement task, we subtracted the mean of all conditions with descending sounds from the mean of all conditions with ascending sounds. A Pearson's Bayesian correlation analysis

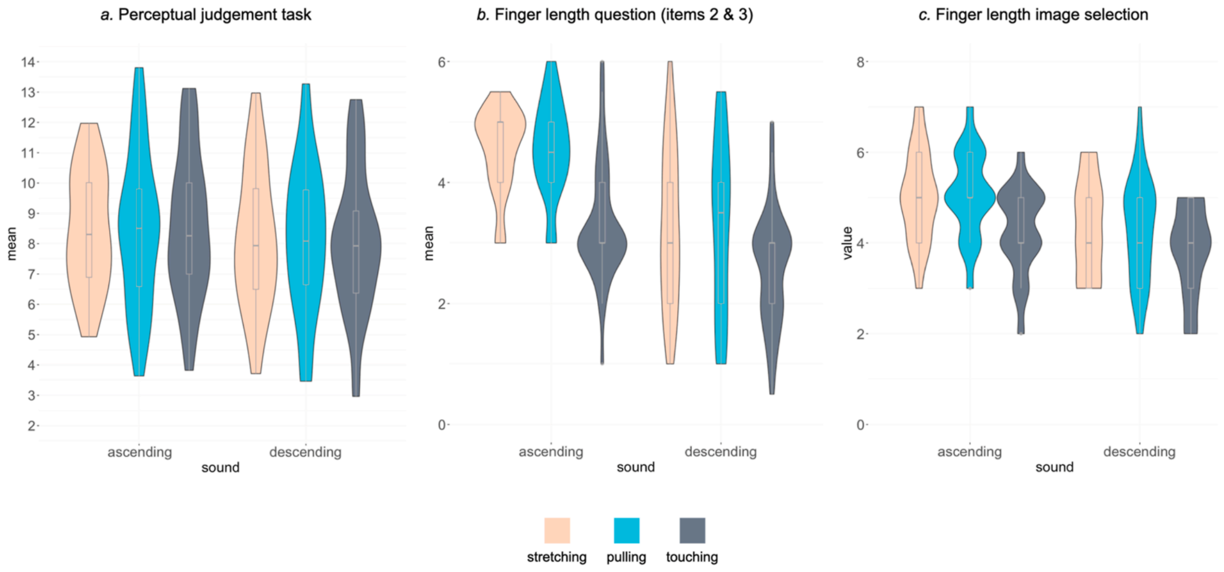

Fig. 5 – Plots showing means, standard errors, central tendencies, and distribution respectively for: (*a*) the mean estimated finger length, (*b*) the mean of questions 2 and 3 (*inverted*) of the finger elongation and embodiment questionnaire, and (*c*) the image best representing participants felt finger size pertaining to Experiment 2. The evaluated factors are action (stretching, pulling, and touching), and sound (ascending, descending).

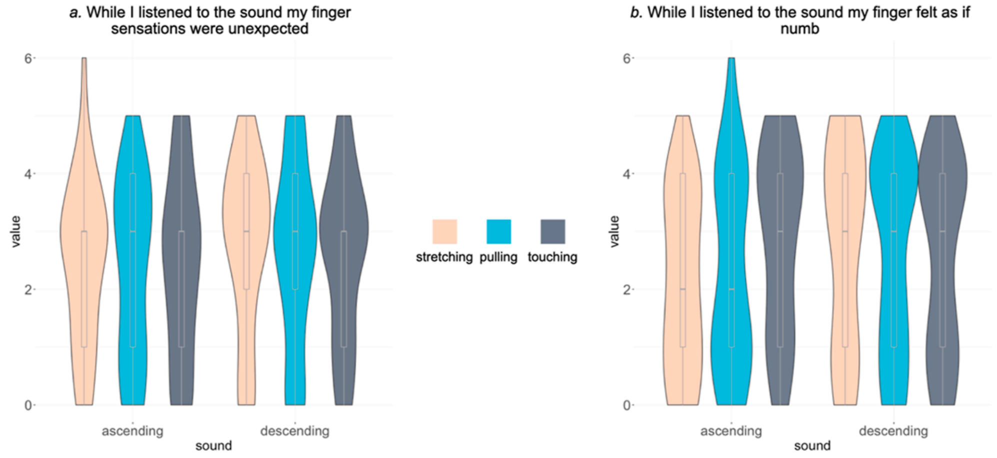

Fig. 6 – Plots showing means, standard errors, central tendencies, and distribution respectively for: (*a*) the feeling that the finger sensations were unexpected, (*b*) the feeling of finger numbness pertaining to Experiment 2. The evaluated factors are action (stretching, pulling, and touching), and sound (ascending, descending).

was performed for this difference and the mean value of the Plymouth sensory imagery questionnaire, bodily, and general imagery. No correlation was found for either auditory ($r = -.118$, $BF_{10} = .230$), bodily ($r = -.067$, $BF_{10} = .225$), nor general imagery ($r = -.065$, $BF_{10} = .229$).

For the questionnaire, as in Experiment 1, we calculated the mean responses for the questionnaire items 2 and inverted item 3 (respectively corresponding to perceiving the finger as longer and shorter). The mean response for all descending conditions was then subtracted from the mean response for all ascending conditions. A correlation analysis was performed between such differences and the same three dimensions imagery questionnaire. Bayesian Kendall's Tau-b tests showed no correlation for either auditory ($\tau_b = -.098$, $BF_{10} = .305$), bodily ($\tau_b = .021$, $BF_{10} = .225$), and general imagery ($\tau_b = -.017$, $BF_{10} = .0224$).

4. Discussion

In this study, we sought to assess how different top-down (i.e., attention, sensory imagery) and bottom-up (i.e., different multisensory couplings) cues contribute to the Auditory Pinocchio illusion (Nava & Tajadura-Jiménez, 2020; Tajadura-Jiménez, Vakali, et al., 2017). We assessed whether they partake in anchoring an arbitrary sound to the finger to produce changes in spatial body perception. In Experiment 1 we showed that attending to the finger contributed to alterations in perceived finger length according to our questionnaire ratings but not according to our perceptual judgement task. Remarkably, answers to item 14 (Table 1) indicated that participants reported associating the sound to the finger more when attending it than when not attending it, confirming a top-down association which did not impact perceptual judgements. In contrast, Experiment 2 showed that the tested multisensory actions (stretching, pulling, or touching) all contributed to changes in finger length perception according to questionnaires and the perceptual judgement task, and that these were in fact correlated. For the touching condition less enlargement was reported in the questionnaires compared to the other conditions, which seems to imply that signal redundancy did not impact their ratings, but the consistency with vertical displacement across modalities (actions and sounds jointly indicative of verticality) did.

In our perceptual judgement task, participants estimated the location of their occluded fingertip and base relying on proprioceptive and somatosensory cues, from which we inferred their estimated finger length. Somatosensory and proprioceptive estimates of one's body metrics have been found to be largely distorted compared to explicit judgements such as images depicting different hand sizes (Longo and Haggard, 2010; Longo 2015). Interestingly, these distortions approximately correspond to early somatosensory maps which represent the body surface according to the density of cortical receptive fields (Penfield & Rasmussen, 1950; Longo 2015), suggesting a link between somatosensory size estimates and cortical characteristics. In fact, transcranial stimulation on the somatosensory cortex results in an increased estimated hand size (Giurgola et al., 2019). By relying on somatosensory and proprioceptive cues, our measure is thus thought to assess judgments reliant on lower-level aspects of body perception, while our questionnaire (including the item depicting various images) is thought to assess higher-level judgments about one's body.

Previous research with the Auditory Pinocchio Illusion found that while adults showed perceptual judgments and questionnaire ratings consistent with the glissando's direction, children did so only in the questionnaires but not in their perceptual ratings (Nava & Tajadura-Jiménez, 2020). The authors claimed that this respectively corresponded to differences in the development of the body schema and body image between children and adults. Such a different pattern between children and adults for top-down and bottom-up judgements has been previously reported in the literature on body perception (Cowie, Makin, & Bremner, 2013; Nava, Bolognini, & Turati, 2017, 2018; but see also Weijs, Roel Lesur, Daum, & Lenggenhager, 2024). In adults, studies involving multisensory manipulations of one's body perception, have shown that explicit ratings are not necessarily consistent with behaviours or physiological responses (Roel Lesur, Weijs, et al., 2020; Rohde et al., 2011; Macauda et al., 2015; Roel Lesur, Stussi, Bertrand, Delplanque, & Lenggenhager, 2023). In our study, while top-down cues contributed to alterations according to questionnaire ratings, perceptual judgments only changed when lower-level somatosensory signals were involved. It seems that top-down conceptual associations are necessary for this illusion to occur, but only together with bottom-up factors do they recalibrate lower-level aspects of perception. The need for a conceptual association is confirmed by previous research showing that the illusion does not occur when performing the task horizontally which is inconsistent with the expected association between pitch and verticality (Nava & Tajadura-Jiménez, 2020). Our study shows that while these semantic associations are necessary, they are not sufficient to impact perceptual judgements, where somatosensory cues are additionally needed.

In Experiment 2, while all multisensory conditions contributed to changes in spatial perception, interesting findings emerged from the touching condition. Compared to stretching or pulling, here participants estimated their finger size as smaller in our questionnaire measures but not according to perceptual judgments. Participants also reported a greater feeling of numbness in the touching condition for ascending sounds, which might be linked to the inconsistency between the expected vertical displacement (driven by the sound) and the non-displacing action of *touching* the finger. Previous research has shown that breaking expectations regarding bodily signals might contribute to feelings of numbness and other symptoms of disembodiment (Roel Lesur, Weijs, et al., 2020; Roel Lesur et al., 2021).

An additional interest of our study was to assess the degree to which sensory imagery might play a role in this alteration of body and spatial perception. However, according to our measures we found no link between the illusion ratings and sensory imagery. A previous study found a correlation between tactile imagery and the susceptibility to a body-weight illusion task (D'Adamo et al., 2024). However such findings are inconsistent across different measures in such a study and did not compare for multiple comparisons. According to our measures, mental imagery traits do not play a role in this illusion.

5. Limitations

Readers should note that in Experiment 1 we used a delay (of 500 ms) between instructing the finger side and playing the sound. This, together with the countdown and 500 ms before the cue, was based on a study showing a similar degree of sensory attenuation for self- and externally generated auditory cues following a countdown and a 500 ms pause before the cue (Kaiser & Schütz-Bosbach, 2018). Sensory attenuation is usually considered a measure of self-generated signals (Kiepe, Kraus, & Hesselmann, 2021). However, the countdown in that study was presented visually and the task and purposes of the study were different. It could be that reducing the interval between attending and playback might give further insights or result in a greater sense of self-generation of the sound. In any case, we found differences when participants reported (in questionnaire ratings) feeling that attending the finger produced the sound between attention conditions (see Fig. 4). Future studies might consider physiological markers related to attention to trigger the sounds.

The questionnaire we used was taken from previous studies using the Auditory Pinocchio illusion, however some of the items may not have been appropriate for our study. For example, results for item 6 suggest that participants felt their finger elongating had a clear effect of action that may be due simply to the fact that participants indeed stretched their finger (and elongated the corresponding muscles) in two of the conditions. However, and importantly, there was an effect of sound that was independent of the action, with finger larger elongation on ascending compared to descending sounds. Furthermore, during debriefing, a few participants reported that they felt numbness due to the positioning and length of the task which may have influenced their responses for the corresponding question (item 13).

6. Conclusions and outlook

Our study replicated the Auditory Pinocchio illusion using a higher resolution and more controlled task to assess perceptual judgements. It revealed that while attention is enough to produce changes in body size estimates according to questionnaires, multisensory signals are necessary to affect perceptual judgements. In this sense it seems that top-down associations between the pitch and verticality are elemental, but multisensory signals are necessary on top of that for affecting lower-level aspects of perception. Furthermore, when it comes to multisensory couplings any of the tested multisensory combinations involving the finger similarly contributed to changes in perceptual judgements. At large, multisensory actions that are consistent with vertical extension seem to play a stronger role in perceived finger length extension according to both questionnaire ratings and perceptual judgements. These findings may influence future developments of auditory prostheses or other applications that need anchoring auditory cues to body parts (Ley-Flores et al., 2022; Singh, et al., 2024). In the context of body transformation experiences, auditory compared to visual cues have the benefit of being more transparent so that they can be used during everyday tasks without much interference. This may be important for applications involving movement while directing vision to other necessary tasks (D'Adamo et al., 2024; Ley-Flores, Bevilacqua, Bianchi-Berthouze, & Taiadura-Jiménez, 2019; Singh et al., 2024).

CRediT authorship contribution statement

Marte Roel Lesur: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Matthew R. Longo:** Writing – review & editing, Conceptualization. **Ana Tajadura-Jiménez:** Writing – review & editing, Resources, Methodology, Funding acquisition, Conceptualization.

Acknowledgements

We would like to thank Joaku de Sotavento for his support developing the first prototype for the finger-length estimation task and for developing the distance sensor. Also, Marina Palomar Pampyn for her support with the anatomical terminology. We further extend our gratitude to the whole i_mBODY Lab for their support. This project was funded by the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement No. 101002711) and the MICIU/AEI (project PID2023-150259OB-C21). The article processing costs were covered by the Universidad Carlos III de Madrid (Agreement CRUE-Madroño 2025). Funding for article processing fees: Universidad Carlos III de Madrid (agreement CRUE-Madroño 2025).

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2025.06.014.

REFERENCES

Abdi, H. (2007). The bonferonni and Šidák corrections for multiple comparisons. In N. Salkind (Ed.), *Encyclopedia of measurement and statistics* (Thousand Oaks).

Andrade, J., May, J., Deeprose, C., Baugh, S.-J., & Ganis, G. (2014). Assessing vividness of mental imagery: The Plymouth sensory imagery questionnaire. *British Journal of Psychology, 105*(4), 547–563. https://doi.org/10.1111/bjop.12050

Apps, M. A. J., & Tsakiris, M. (2014). The free-energy self: A predictive coding account of self-recognition. *Neuroscience and Biobehavioral Reviews, 41*, 85–97. https://doi.org/10.1016/j.neubiorev.2013.01.029

Bertelson, P., & Aschersleben, G. (1998). Automatic visual bias of perceived auditory location. *Psychonomic Bulletin & Review, 5*(3), 482–489.

Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience, 13*(8), 556–571. https://doi.org/10.1038/nrn3292

Bolt, E., Ho, J. T., Roel Lesur, M., Soutschek, A., Tobler, P. N., & Lenggenhager, B. (2021). Effects of a virtual gender swap on social and temporal decision-making. *Scientific Reports, 11*(1), Article 15376. https://doi.org/10.1038/s41598-021-94869-z

Botvinick, M., & Cohen, J. (1998). Rubber hands "feel" touch that eyes see. *Nature, 391*(6669), 756. https://doi.org/10.1038/35784

Caclin, A., Soto-Faraco, S., Kingstone, A., & Spence, C. (2002). Tactile "capture" of audition. *Perception & Psychophysics, 64*(4), 616–630.

Cowie, D., Makin, T. R., & Bremner, A. J. (2013). Children's responses to the rubber-hand illusion reveal dissociable pathways in body representation. *Psychological Science, 24*(5), 762–769. https://doi.org/10.1177/0956797612462902

D'Adamo, A., Roel Lesur, M., Turmo Vidal, L., Dehshibi, M. M., De La Prida, D., Diaz-Durán, J. R., et al. (2024). SoniWeight shoes: Investigating effects and personalization of a wearable sound device for altering body perception and behavior. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–20. https://doi.org/10.1145/3613904.3642651

De Vignemont, F. (2017). *Mind the body* (Vol. 1). Oxford University Press. https://doi.org/10.1093/oso/9780198735885.001.0001

de Vignemont, F., Ehrsson, H. H., & Haggard, P. (2005). Bodily illusions modulate tactile perception. *Current Biology, 15*(14), 1286–1290. https://doi.org/10.1016/j.cub.2005.06.067

Deroy, O., Fernandez-Prieto, I., Navarra, J., & Spence, C. (2018). Unraveling the paradox of spatial pitch. In T. L. Hubbard (Ed.), *Spatial biases in perception and cognition* (1st ed., pp. 77–93). Cambridge University Press. https://doi.org/10.1017/9781316651247.006.

Dienes, Z., Lush, P., Palfi, B., Roseboom, W., Scott, R., Parris, B., et al. (2022). Phenomenological control as cold control. *Psychology of Consciousness: Theory, Research, and Practice, 9*(2), 101–116. https://doi.org/10.1037/cns0000230

Dijkerman, C., & Lenggenhager, B. (2018). The body and cognition: The relation between body representations and higher level cognitive and social processes. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 104*, 133–139. https://doi.org/10.1016/j.cortex.2018.06.001

Ehrsson, H. H. (2012). The concept of body ownership and its relation to multisensory integration. In B. E. Stein (Ed.), *The new handbook of multisensory processes, B.E. stein* (pp. 775–792). Cambridge, MA, USA: MIT Press, 2012.

Ehrsson, H. H., Fotopoulou, A., Radziun, D., Longo, M. R., & Tsakiris, M. (2022). No specific relationship between hypnotic suggestibility and the rubber hand illusion. *Nature Communications, 13*(1). https://doi.org/10.1038/s41467-022-28177-z. Article 1.

Eitan, Z., Schupak, A., Gotler, A., & Marks, L. E. (2014). Lower pitch is larger, yet falling pitches shrink. *Experimental Psychology*. https://econtent.hogrefe.com/doi/10.1027/1618-3169/a000246.

Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision, 10*(1), 6. https://doi.org/10.1167/10.1.6

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

Giurgola, S., Pisoni, A., Maravita, A., Vallar, G., & Bolognini, N. (2019). Somatosensory cortical representation of the body size. *Human Brain Mapping, 40*(12), 3534–3547. https://doi.org/10.1002/hbm.24614

Hoort van der, B., Guterstam, A., & Ehrsson, H. H. (2011). Being barbie: The size of one's own body determines the perceived size of the world. *Plos One, 6*(5), Article e20195. https://doi.org/10.1371/journal.pone.0020195

Kaiser, J., & Schütz-Bosbach, S. (2018). Sensory attenuation of self-produced signals does not rely on self-specific motor predictions. *European Journal of Neuroscience, 47*(11), 1303–1310. https://doi.org/10.1111/ejn.13931

Kammers, M. P. M., Longo, M. R., Tsakiris, M., Dijkerman, H. C., & Haggard, P. (2009). Specificity and coherence of body representations. *Perception, 38*(12), 1804–1820. https://doi.org/10.1068/p6389

Kay, M., & Wobbrock, J. (2020). *ARTool: Aligned rank transformation for nonparametric factorial ANOVAs*. https://doi.org/10.5281/zenodo.594511 (Version R package version 0.10.8) [Computer software].

Kiepe, F., Kraus, N., & Hesselmann, G. (2021). Sensory attenuation in the auditory modality as a window into predictive processing. *Frontiers in Human Neuroscience, 15*. https://www.frontiersin.org/articles/10.3389/fnhum.2021.704668.

Kilteni, K., Maselli, A., Kording, K. P., & Slater, M. (2015). Over my fake body: Body ownership illusions for studying the multisensory basis of own-body perception. *Frontiers in Human Neuroscience, 9*. https://doi.org/10.3389/fnhum.2015.00141

Lackner, J. R. (1988). Some proprioceptive influences on the perceptual representation of body shape and orientation. *Brain: a Journal of Neurology, 111*(2), 281–297. https://doi.org/10.1093/brain/111.2.281

Lenggenhager, B., Mouthon, M., & Blanke, O. (2009). Spatial aspects of bodily self-consciousness. *Consciousness and Cognition, 18*(1), 110–117. https://doi.org/10.1016/j.concog.2008.11.003

Ley-Flores, J., Alshami, E., Singh, A., Bevilacqua, F., Bianchi-Berthouze, N., Deroy, O., et al. (2022). Effects of pitch and musical sounds on body-representations when moving with sound. *Scientific Reports, 12*(1). https://doi.org/10.1038/s41598-022-06210-x. Article 1.

Ley-Flores, J., Bevilacqua, F., Bianchi-Berthouze, N., & Taiadura-Jiménez, A. (2019). Altering body perception and emotion in physically inactive people through movement sonification. In *2019 8th international conference on affective computing and intelligent interaction (ACII)* (pp. 1–7). https://doi.org/10.1109/ACII.2019.8925432

Lush, P., Botan, V., Scott, R. B., Seth, A., Ward, J., & Dienes, Z. (2019). *Phenomenological control: Response to imaginative suggestion predicts measures of mirror touch synaesthesia, vicarious pain and the rubber hand illusion*. https://doi.org/10.31234/osf.io/82jav

Lush, P., Botan, V., Scott, R. B., Seth, A. K., Ward, J., & Dienes, Z. (2020). Trait phenomenological control predicts experience of mirror synaesthesia and the rubber hand illusion. *Nature Communications, 11*(1). https://doi.org/10.1038/s41467-020-18591-6. Article 1.

Macauda, G., Bertolini, G., Palla, A., Straumann, D., Brugger, P., & Lenggenhager, B. (2015). Binding body and self in visuo-vestibular conflicts. *The European Journal of Neuroscience, 41*(6), 810–817. https://doi.org/10.1111/ejn.12809

Maister, L., Slater, M., Sanchez-Vives, M. V., & Tsakiris, M. (2015). Changing bodies changes minds: Owning another body affects social cognition. *Trends in Cognitive Sciences, 19*(1), 6–12. https://doi.org/10.1016/j.tics.2014.11.001

Mancini, F., Longo, M. R., Kammers, M. P. M., & Haggard, P. (2011). Visual distortion of body size modulates pain perception. *Psychological Science, 22*(3), 325–330. https://doi.org/10.1177/0956797611398496

Marino, B. F. M., Stucchi, N., Nava, E., Haggard, P., & Maravita, A. (2010). Distorting the visual size of the hand affects hand pre-shaping during grasping. *Experimental Brain Research, 202*(2), 499–505. https://doi.org/10.1007/s00221-009-2143-4

Nava, E., Bolognini, N., & Turati, C. (2017). The development of a cross-modal sense of body ownership. *Psychological Science, 28*(3), 330–337. https://doi.org/10.1177/0956797616682464

Nava, E., Gamberini, C., Berardis, A., & Bolognini, N. (2018). Action shapes the sense of body ownership across human development. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.02507

Nava, E., & Tajadura-Jiménez, A. (2020). Auditory-induced body distortions in children and adults. *Scientific Reports, 10*(1). https://doi.org/10.1038/s41598-020-59979-0. Article 1.

Navas-León, S., Morales Márquez, L., Sánchez-Martín, M., Crucianelli, L., Bianchi-Berthouze, N., Borda-Mas, M., et al. (2023). Exploring multisensory integration of non-naturalistic sounds on body perception in young females with eating disorders symptomatology: A study protocol. *Journal of Eating Disorders, 11*(1), 28. https://doi.org/10.1186/s40337-023-00749-4

Newbold, J., Gold, N. E., & Bianchi-Berthouze, N. (2020). Movement sonification expectancy model: Leveraging musical expectancy theory to create movement-altering sonifications. *Journal on Multimodal User Interfaces, 14*(2), 153–166. https://doi.org/10.1007/s12193-020-00322-2

Pérez-Fabello, M. J., & Campos, A. (2020). Spanish version of the Plymouth sensory imagery questionnaire. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.00916

Palmer, R. D., & Field, P. B. (1968). Visual imagery and susceptibility to hypnosis. *Journal of Consulting and Clinical Psychology, 32*(4), 456–461. https://doi.org/10.1037/h0026053

Parise, C. V., Knorre, K., & Ernst, M. O. (2014). Natural auditory scene statistics shapes human spatial hearing. *Proceedings of the National Academy of Sciences, 111*(16), 6104–6108. https://doi.org/10.1073/pnas.1322705111

Pavani, F., & Zampini, M. (2007). The role of hand size in the fake-hand illusion paradigm. *Perception, 36*(10), 1547–1554. https://doi.org/10.1068/p5853

Riemer, M., Trojan, J., Beauchamp, M., & Fuchs, X. (2019). The rubber hand universe: On the impact of methodological differences in the rubber hand illusion. *Neuroscience and Biobehavioral Reviews,* 104, 268–280. https://doi.org/10.1016/j.neubiorev.2019.07.008

Roel Lesur, M., Aicher, H., Delplanque, S., & Lenggenhager, B. (2020a). Being short, sweet, and sour: Congruent visuo-olfactory stimulation enhances illusory embodiment. *Perception,* 49(6), 693–696. https://doi.org/10.1177/0301006620928669

Roel Lesur, M., Bolt, E., Saetta, G., & Lenggenhager, B. (2021). The monologue of the double: Allocentric reduplication of the own voice alters bodily self-perception. *Consciousness and Cognition,* 95, Article 103223. https://doi.org/10.1016/j.concog.2021.103223

Roel Lesur, M., Gaebler, M., Bertrand, P., & Lenggenhager, B. (2018). The plasticity of the bodily self: Head movements in bodily illusions and their relation to Gallagher's body image and body schema. *Constructivist Foundations, 14*(1), 94–105.

Roel Lesur, M., Stussi, Y., Bertrand, P., Delplanque, S., & Lenggenhager, B. (2023). Different armpits under my new nose: Olfactory sex but not gender affects implicit measures of embodiment. *Biological Psychology, 176*, Article 108477. https://doi.org/10.1016/j.biopsycho.2022.108477

Roel Lesur, M., Weijs, M. L., Simon, C., Kannape, O. A., & Lenggenhager, B. (2020b). Psychometrics of disembodiment and its differential modulation by visuomotor and visuotactile mismatches. *iScience, 23*(3), Article 100901. https://doi.org/10.1016/j.isci.2020.100901

Rohde, M., Luca, M. D., & Ernst, M. O. (2011). The rubber hand illusion: Feeling of ownership and proprioceptive drift do not Go hand in hand. *Plos One,* 6(6), Article e21659. https://doi.org/10.1371/journal.pone.0021659

Singh, A., Hrobat, M., Gui, S., Bianchi-Berthouze, N., Ley-Flores, J., Bevilacqua, F., et al. (2024). Pushed by sound: Effects of sound and movement direction on body perception, experience quality, and exercise support. *ACM Trans. Comput.-Hum. Interact., 31*(4). https://doi.org/10.1145/3648616, 53:1-53:36.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics, 73*(4), 971–995. https://doi.org/10.3758/s13414-010-0073-7

Spence, C. (2019). On the relative nature of (Pitch-Based) crossmodal correspondences. *Multisensory Research, 32*(3), 235–265. https://doi.org/10.1163/22134808-20191407

Tajadura-Jiménez, A., Banakou, D., Bianchi-Berthouze, N., & Slater, M. (2017a). Embodiment in a child-like talking virtual body influences object size perception, self-identification, and subsequent real speaking. *Scientific Reports, 7*(1), 9637. https://doi.org/10.1038/s41598-017-09497-3

Tajadura-Jiménez, A., Deroy, O., Marquardt, T., Bianchi-Berthouze, N., Asai, T., Kimura, T., et al. (2018). Audio-tactile cues from an object's fall change estimates of one's body height. *Plos One, 13*(6), Article e0199354. https://doi.org/10.1371/journal.pone.0199354

Tajadura-Jiménez, A., Fairhurst, M. T., & Deroy, O. (2022). Sensing the body through sound. In A. J. T. Alsmith, & M. R. Longo (Eds.), *The routledge handbook of bodily awareness* (1st ed., pp. 230–246). Routledge. https://doi.org/10.4324/9780429321542-21.

Tajadura-Jiménez, A., Väljamäe, A., Bevilacqua, F., & Bianchi-Berthouze, N. (2017c). Principles for designing body-centered auditory feedback. In K. L. Norman, & J. Kirakowski (Eds.), *The wiley handbook of human computer interaction* (pp. 371–403). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118976005.ch18.

Tajadura-Jiménez, A., Väljamäe, A., Toshima, I., Kimura, T., Tsakiris, M., & Kitagawa, N. (2012). Action sounds recalibrate perceived tactile distance. *Current Biology, 22*(13), R516–R517. https://doi.org/10.1016/j.cub.2012.04.028

Tajadura-Jiménez, A., Vakali, M., Fairhurst, M. T., Mandrigin, A., Bianchi-Berthouze, N., & Deroy, O. (2017b). Contingent sounds change the mental representation of one's finger length. *Scientific Reports, 7*(1). https://doi.org/10.1038/s41598-017-05870-4. Article 1.

Tsakiris, M., & Haggard, P. (2005). The rubber hand illusion revisited: Visuotactile integration and self-attribution. *Journal of Experimental Psychology. Human Perception and Performance, 31* (1), 80–91. https://doi.org/10.1037/0096-1523.31.1.80

Van Den Bergh, D., Van Doorn, J., Marsman, M., Draws, T., Van Kesteren, E.-J., Derks, K., et al. (2020). A tutorial on conducting and interpreting a bayesian ANOVA in JASP. *L'Année Psychologique, 120*(1), 73–96. https://doi.org/10.3917/anpsy1.201.0073

van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., et al. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review, 28*(3), 813–826. https://doi.org/10.3758/s13423-020-01798-5

Vroomen, J., & De Gelder, B. (2004). Perceptual effects of crossmodal stimulation: The cases of ventriloquism and the freezing phenomenon. In G. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of multisensory processes* (pp. 141–150). Cambridge: MIT Press.

Walker, L., & Walker, P. (2016). Cross-sensory mapping of feature values in the size–brightness correspondence can be more relative than absolute. *Journal of Experimental Psychology. Human Perception and Performance, 42*(1), 138–150. https://doi.org/10.1037/xhp0000128

Weijs, M. L., Roel Lesur, M., Daum, M. M., & Lenggenhager, B. (2024). Keeping up with ourselves: Multimodal processes underlying body ownership across the lifespan. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 177*, 209–223. https://doi.org/10.1016/j.cortex.2024.05.013

Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011). The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 143–146). https://doi.org/10.1145/1978942.1978963