



**HAL**  
open science

## **Ears on the Hand: Reaching Three-Dimensional Targets With an Audio-Motor Device**

Sylvain Hanneton, Thomas Hoellinger, Vincent Forma, Agnes Roby-Brami,  
Malika Auvray

► **To cite this version:**

Sylvain Hanneton, Thomas Hoellinger, Vincent Forma, Agnes Roby-Brami, Malika Auvray. Ears on the Hand: Reaching Three-Dimensional Targets With an Audio-Motor Device. *Multisensory Research*, 2020, 4-5, pp.433-455. 10.1163/22134808-20191436 . hal-02475383

**HAL Id: hal-02475383**

**<https://hal.science/hal-02475383>**

Submitted on 31 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1  
2  
3  
4 **Ears on the hand.**

5  
6 **Reaching three-dimensional targets with an audio-motor device**

7  
8  
9  
10 Sylvain Hanneton<sup>1</sup>, Thomas Hoellinger<sup>2</sup>, Vincent Forma<sup>3</sup>,  
11 Agnes Roby-Brami<sup>4,5</sup>, & Malika Auvray<sup>4</sup>  
12  
13  
14

15 1. Institut des Sciences du Sport-Santé EA3625, Université Paris Descartes, Paris, France

16  
17 2. Laboratoire de Neurophysiologie et Biomécanique du mouvement, Université Libre de  
18 Bruxelles, Faculté des Sciences de la motricité, Brussel, Belgium

19  
20 3. Laboratoire Psychologie de la Perception, CNRS UMR 8242, Université Paris  
21 Descartes, Paris, France

22  
23  
24 4. Institut des Systèmes Intelligents et de Robotique, ISIR, CNRS UMR 7222, Sorbonne  
25 Université, Paris, France.

26  
27 5. Institut des Systèmes Intelligents et de Robotique, Equipe Agathe, INSERM U 1150.  
28  
29  
30

31  
32 **Corresponding author**

33 Malika Auvray, Institut des Systèmes Intelligents et de Robotique, Université Pierre et Marie  
34 Curie, 4 place Jussieu, 75005 Paris, France

35  
36 E-mail: [auvray@isir.upmc.fr](mailto:auvray@isir.upmc.fr)  
37  
38  
39

40  
41 **Short title: Ears on the hand**  
42  
43

44  
45 **Re-submitted to: *Multisensory Research* (June 2019)**  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 **Abstract**  
5  
6  
7

8 Understanding the processes underlying sensorimotor coupling with the environment is  
9 crucial for sensorimotor rehabilitation and sensory substitution. In doing so, devices which  
10 provide novel sensory feedback consequent to body movement may be optimized in order  
11 to enhance motor performance for particular tasks. The aim of the study reported here was  
12 to investigate audio-motor coupling when the auditory experience is linked to movements  
13 of the head or the hands. The participants had to localize and reach a virtual source with  
14 the dominant hand in response to sounds. An electromagnetic system recorded the position  
15 and orientation of the participants' head and hands. This system was connected to a 3D  
16 audio system that provided binaural auditory feedback on the position of the virtual  
17 listener located on the participants' body. The listener's position was computed either  
18 from the hands or from the head. For the hand condition, the virtual listener was placed on  
19 the dominant hand (the one used to reach the target) in Experiment 1 and on the non-  
20 dominant hand, which was constrained in order to have similar amplitude and degrees of  
21 freedom as that of the head, in Experiment 2. The results revealed that, in the two  
22 experiments, the participants were able to localize a source within the 3D auditory  
23 environment. Performance varied as a function of the effector's degrees of freedom and  
24 the spatial coincidence between sensor and effector. The results also allowed  
25 characterizing the kinematics of the hand and head and how they change with audio-motor  
26 coupling condition and practice.  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43

44 **Keywords:** Sensory substitution, neural plasticity, perceptual adaptation, rehabilitation  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 **Introduction**

5  
6 *Sensory substitution*

7  
8 In the field of perceptual rehabilitation, sensory substitution devices were developed  
9 with the aim of compensating one or several functions of a deficient sensory modality. To do  
10 so, these devices convert stimuli that are normally accessed through this deficient sensory  
11 modality into stimuli accessible by another sensory modality (e.g., Auvray and Myin, 2008;  
12 Gallay *et al.*, 2012, for reviews). For instance, these devices can convert visual information  
13 into tactile stimuli (e.g., Bach-y-Rita *et al.*, 1969) or sounds (e.g., Hanneton *et al.*, 2010;  
14 Meijer, 1992). Studies conducted on sensory substitution revealed an important plasticity of  
15 the central nervous system. Behavioural studies revealed that sensory substitution devices  
16 allow improving spatial processing in the absence of vision. In particular, users of these  
17 systems become able to perform localisation tasks (e.g., Levy-Tzedek *et al.*, 2012; Proulx *et*  
18 *al.*, 2008), recognition tasks (Arno *et al.*, 2001; Auvray *et al.*, 2007-a, 2007-b; Pollok *et al.*,  
19 2005; Sampaio *et al.*, 2001), navigation tasks (Chebat *et al.*, 2011; 2015) and, to some extent,  
20 consolidate a perceptual space that possesses visual characteristics (e.g., Renier *et al.*, 2005).  
21 However, the degree to which the use of a sensory substitution device involves visual  
22 processes varies from one person to another (Arnold *et al.*, 2017). In parallel, brain-imaging  
23 studies have shown that the use and practice of visual-to-tactile (Ptito *et al.*, 2005) and  
24 visual-to-auditory (e.g., Striem-Amit and Amedi, 2014) devices result in increased activation  
25 in blind people's visual cortex (for reviews, see Proulx *et al.*, 2014; Ptito *et al.*, 2018).  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41

42 *Sensory substitution, learning, and body kinematics*

43  
44 Effective use of sensory substitution devices is largely dependent upon the users'  
45 ability to master their different features and the amount of training required to do so. In most  
46 cases, users are able to approximate objects' location and shape in a short amount of time  
47 (i.e., one or two hours with the Tongue Display Unit, Kaczmarek and Haase, 2003). However,  
48 to gain precision in these tasks and to perform more complex ones, more intensive training is  
49 required. The length of the training is estimated at around 10 to 15 hours with visual-to-  
50 auditory devices (Auvray *et al.*, 2007) and around eight hours with visual-to-tactile devices  
51 (Kaczmarek and Haase, 2003). It is estimated that further time still would be required to  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 achieve a high level of expertise, for instance to be able to recognize facial expressions with  
5 visual-to-auditory devices (Striem-Amit *et al.*, 2012).  
6  
7

8  
9 Another crucial characteristic of learning is the ability to generalize, that is, the  
10 ability to extend the acquired perceptual abilities to both new stimuli and new perceptual  
11 conditions. With tactile devices, it has been shown that learning transfers across body  
12 surfaces, both for surfaces represented in adjacent and non-adjacent areas of the  
13 somatosensory cortex (Arnold and Auvray, 2014). This result provides behavioural support  
14 for the claim that training in sensory substitution results in a relative independence from the  
15 stimulated body surface. In addition, transfer of learning to novel stimuli has been  
16 documented for both tactile (Arnold and Auvray, 2018; Kim and Zatorre, 2010) and auditory  
17 devices (Kim and Zatorre, 2008). This ability to generalize to new symbols depends on the  
18 number of stimuli to learn and on feature variability (Arnold and Auvray, 2018). These  
19 results are in line with the reverse hierarchy theory (RHT) of perceptual learning (Ahissar  
20 and Hochstein, 2004), according to which the degree of specificity depends on the difficulty  
21 of the perceptual task and on the level of cortical processing that is required to perform the  
22 task.  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33

34  
35 Finally, performance and experience with sensory substitution devices depend on  
36 each person's abilities (see Arnold *et al.*, 2017, for a review). For instance, Brown *et al.*'s  
37 (2011) study showed that people with good musical abilities learned a visual-to-auditory  
38 device faster than people with low musical abilities, although they did not reach better  
39 abilities with the device. Haigh *et al.*'s (2013) study found that musicians had a better  
40 "visual" acuity (tested with Snellen E) when using the visual-to-auditory conversion system  
41 The vOICE than non-musicians. Note that regarding visual abilities, Brown *et al.* (2011)  
42 found no influence of individual visual imagery abilities (measured with VVIQ, Marks,  
43 1973) on the learning of the vOICE. On the other hand, in a study by Hanneton *et al.* (2015),  
44 performance reached by the participants on a task based on the auditory conversion of a  
45 visual video game was found to be correlated to the scores the participants obtained in a  
46 visual mental rotation test (Vandenberg, 1971). It thus seems that performance with sensory  
47 substitution devices is related to only some visuo-spatial abilities. Overall, these results  
48 suggest that perceptual strategies will differ as a function of people's specificities and  
49 cognitive style which, in some tasks, affects people's performance.  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 The studies described above showed people’s ability to learn how to use sensory  
5 substitution devices. These studies focused on the sensory aspects, leaving the motor changes  
6 involved in learning undocumented. With other devices, learning the new sensory  
7 consequences of body movements has been shown to increase motor performance and to  
8 induce an evolution of body movement characteristics. In particular, as adaptation  
9 progresses, movements become less segmented, smoother, more rapid and fluid, with an  
10 increase in the amplitude of joint motion (e.g., Novak *et al.*, 2003; Rand *et al.*, 2004). Our  
11 study aimed at investigating such motor consequences of perceptual learning in the case of  
12 sensory substitution.  
13  
14  
15  
16  
17  
18  
19  
20  
21

### 22 *Aim of the study*

23  
24 The first aim of the study reported here was to investigate the changes in kinematic  
25 parameters while learning a novel auditory device. To do so, the participants had to learn to  
26 intercept a static auditory target within a 3D immersive auditory environment. Motion  
27 capture sensors were placed on the participants’ head and hand, and their kinematics were  
28 analysed while adapting to this new system. In addition to success rate and reaching duration,  
29 the kinematics of hand and head movements, and in particular the number of velocity peaks,  
30 the direct distance followed by the hand and head during reaching, the cumulative length of  
31 the trajectories, and the cumulative rotations were analysed.  
32  
33  
34  
35  
36  
37  
38

39 The second aim of the study was to investigate if performance with the device is  
40 influenced by the location of the sensors. In particular, one hypothesis raised by sensorimotor  
41 theories of perception is that the use of a head-mounted camera shares more sensorimotor  
42 resemblance with vision than the use of a hand-held camera (e.g., O’Regan, 2011; O’Regan  
43 and Noë, 2001). As a consequence, according to this theory, the use of a head-mounted  
44 camera might favour an association between perception with the device and the visual  
45 sensory modality as compared with the situations in which the camera is hand-held. Thus,  
46 two modes were investigated in our experiment, the head mode and the hand mode. In the  
47 head mode, the head sensor was used to drive the position and orientation of the virtual  
48 listener in the 3D auditory display. In the hand mode, the position and orientation of the  
49 virtual listener was driven by the sensor fixed on the hand. Consequently, in this condition,  
50 the target’s sound is perceived through “ears” on the participant’s hand.  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 Thirdly, in order to further investigate the influence of body kinematics on the use of  
5 the device, our experiments studied the extent to which the differences in performance  
6 obtained with the hand and head is attributable to differences in amplitude of movement and  
7 degrees of freedom. Thus, in the first experiment, hand movements were not constrained,  
8 whereas in the second experiment, the movements of the non-dominant hand wearing the  
9 listener were constrained in order to have similar amplitude and degrees of freedom as that of  
10 the head.  
11  
12  
13  
14  
15  
16  
17  
18

## 19 **Experiment 1**

### 20 **Materials and methods**

#### 21 *Participants*

22  
23  
24 12 participants (4 females and 8 males, age range: 25-50 years) took part in this  
25 experiment which lasted approximately 40 minutes. All were students or faculty members  
26 at Paris Descartes University. Inclusion criteria were being aged between 18 and 50 years  
27 with no self-reported neurological disorder nor auditory impairment. The participants were  
28 naive as to the hypotheses under investigation. The experiment was performed in  
29 accordance with the ethical standards laid down in the 1991 Declaration of Helsinki.  
30  
31  
32  
33  
34  
35  
36

#### 37 *Experimental set-up*

38  
39 An electromagnetic device (Fastrak Polhemus) was connected to the 3D audio  
40 system OpenAL that provides auditory feedback on movements. The Polhemus system  
41 acquires 6D motion data with a 30 Hz sample frequency. This electromagnetic system  
42 gives real time position (X, Y, Z) and orientation (Euler angles: azimuth, elevation, and  
43 roll) of two sensors, allowing a distinction between changes in position and orientation.  
44 The body-auditory coupling system is thus made of a listener (sensor) placed on the  
45 participants' body and a virtual audio source. The perceived sounds depend on the position  
46 and orientation of the listener relative to the position of the source (see Figure 1). More  
47 precisely, a «buzzing» sound similar to a fly with significant harmonics between 100 and  
48 2000 Hz varied in intensity and frequency as a function of the position and orientation of  
49 the listener (see below for the modelling and equations used).  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 In the “head mode”, the listener was attached to the head and the perceived sound  
5 was sensitive to the position and orientation of the head with respect to the sound source.  
6 It corresponded to a relatively natural situation since the positions of the “virtual ears”  
7 were located close to the real ears. In the “hand mode”, the listener was attached to the  
8 hand used to reach the fly and the perceived sound was sensitive to the position and  
9 orientation of the hand relative to the presented sound source (see figure 1).  
10  
11  
12  
13  
14

15 (Fig. 1 about here)  
16  
17  
18

### 19 *Sound spatialisation modelling*

20 The binaural sound generated by the audio device (SoundMax integrated digital  
21 audio by Analog Devices Inc.) was computed from the relative position and orientation  
22 between the listener and the virtual source (see figure 1). The 3D audio system was based  
23 on the library OpenAL<sup>1</sup>, used in video games, which allows 3D sound spatialisation.  
24 OpenAL uses both interaural intensity differences (IIDs) and interaural time differences  
25 (ITDs), which are combined to code for the target’s distance and direction. As a sound  
26 source does not have the same distance between the two ears, the sound arrives faster to  
27 the closest ear (ITDs) and with a higher intensity (IIDs). Regarding IIDs, according to the  
28 specifications of OpenAL library, the intensity decreases and the sound is slightly muffled  
29 as the listener moves away from the source. This models reality in which air absorbs high  
30 frequency waves more rapidly than low frequency ones. Within 5 cm from the source, the  
31 sound remains at a constant maximum volume. Beyond a maximal distance of 40 cm, the  
32 amplitude is held constant. Between these two limits, the intensity of the sound is  
33 computed as a function of distance, for each hear, according to the following equation:  
34  $G(d)=D_{ref}/(D_{ref}+R*(d-D_{ref}))$ . where  $d$  is the distance to the source,  $G$  is the gain applied to  
35 the sound,  $R$  ( $= 1.3$ ) is the Roll-off factor, and  $D_{ref}$  ( $= 5$  cm) is a reference distance  
36 corresponding to the distance that gives a unitary gain.  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53

### 54 *Procedure*

55 For each of the two modes (head and hand), the participants were instructed to catch  
56 the fly, i.e. to reach the static virtual sound source, with their dominant hand. The  
57 participants sat in front of a table. They were blindfolded, fitted with headphones, and two  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 sensors were fixed with specific double-sided tape on the back of the hand and on the  
5 headphones. A geometrical transformation was applied in order to place the virtual listener  
6 either on the centre of the hand or on the centre of the head, midway between the ears. At  
7 the start of each trial, the dominant hand, used to reach the source, was placed on the  
8 sternum near the xyphoid process. The participants were instructed to start to move as  
9 soon as they heard the sound. The trial was considered as being successful (i.e., the “fly”  
10 was caught) when the centre of the hand approached within a distance smaller than 5 cm  
11 of the static sound source. In this case, the recording stopped and a synthetic applause was  
12 played. The participants were able to give their responses up to 31s after the end of the  
13 stimulus.  
14  
15  
16  
17  
18  
19  
20  
21

22 The virtual source could be presented at 9 possible locations within the  
23 participants’ workspace: 1 in the middle and 8 in the corners of a virtual parallelepiped  
24 rectangle (20 cm on the right or the left of the participants’ midline, at 5 or 25 cm from of  
25 their sternum and at 15 or 35 cm above the table). The order of presentation of the 9  
26 targets was randomized within blocks. Blocks of 9 trials were repeated three times for  
27 each of the two modes, with the participants completing 54 trials in total. Half of the  
28 participants began with the hand mode while the other half began with the head mode.  
29  
30  
31  
32  
33  
34  
35  
36

### 37 *Statistical analyses*

38  
39 The computed dependent variables were the success rate (percentage of catches  
40 before cut-off) and reaching duration (time to reach the target). The kinematics of hand  
41 and head movements during successful trials were characterized by the number of velocity  
42 peaks, the direct distance followed by the hand and head during reaching (3D distance  
43 between their starting and final positions), each sensor’s cumulative length of trajectories  
44 and cumulative rotations. The cumulative length of the trajectories corresponds to the sum  
45 of all the elementary 3D displacements performed during each 33 ms time sample. The  
46 cumulative rotations were obtained after calculating the 3D rotation matrix at each time  
47 sample, then the change in 3D orientation during one time sample. The rotation angle  
48 performed during each time sample was computed using the axis-angle formalism. Then,  
49 the obtained elementary rotations were summed during the whole trial to obtain the  
50 cumulative rotations of the hand and head.  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 The distribution of the variables was verified with Mauchley's Sphericity test.  
5  
6 Repeated measures ANOVAs were conducted to analyse the dependent variables with the  
7  
8 factors Session (2 levels), Mode (2 levels: hand and head), and Block (3 levels). The  
9  
10 results were averaged across targets. Post Hoc comparisons were performed with  
11  
12 Bonferroni tests. A Multivariate approach of ANOVA was used when the conditions for  
13  
14 ANOVA were not fully met, completed with Friedman non-parametric analysis, when  
15  
16 relevant. Additional statistics were performed using Chi-squared test.  
17

## 18 **Results**

### 19 *Success rate*

20  
21 The success rate appeared to be variable across participants and conditions. The  
22  
23 mean success rate in the three blocks ranged between 51.8 and 100% in the hand mode  
24  
25 and between 26 and 100% in the head mode. The success rate did not vary with the Order  
26  
27 of sessions and there was no interaction with Mode (Pearson's Chi-squared test,  $p=0.18$   
28  
29 and  $=0.38$ , respectively). Given this, results were pooled across sessions for the subsequent  
30  
31 analyses.  
32

33 An ANOVA conducted with the factors Mode and Block showed a significant effect  
34  
35 of Mode ( $F[1,11]=15.03$ ,  $p=0.002$ ). The success rate was higher in the hand mode ( $75.93 \pm$   
36  
37  $4.79$  %, mean  $\pm$  sem) than in the head mode ( $71.3 \pm 7.58$  %) (Figure 2A). The sphericity  
38  
39 condition was not fully met for the Block factor (Mauchley's test  $p=0.044$ ). Non  
40  
41 parametric statistics (Friedman test) showed a significant effect of Block in the head mode  
42  
43 ( $p= 0.011$ ) but not in the hand mode ( $p=0.30$ ).  
44

45 (Fig. 2 about here)

### 46 47 *Trial duration*

48 An ANOVA conducted with the Factors Mode and Block on trial duration did not  
49  
50 show any significant main effect, nor interaction between Mode and Block (all  $ps>.05$ ).  
51  
52 Numerically, trial duration was  $11.59 \pm 1.03s$  in the hand mode and  $12.37 \pm 1.02s$  in the  
53  
54 head mode (see Figure 2B).  
55  
56

### 57 58 *Hand and head kinematics*

1  
2  
3  
4 An ANOVA conducted with the Factors Mode and Block on the direct distance  
5 followed by the hand during reaching did not show any effect of Mode nor Block ( $F_s < 1$ ),  
6 confirming that movement execution during successful trials was reproducible across  
7 conditions (see Figure 3A). The trajectories followed by the head and hand did not follow  
8 a single path, but several of them, suggesting a behaviour of exploration. As a  
9 consequence, cumulative trajectories and cumulative rotations of the head and hand were  
10 used as dependent variables. Figure 3 shows the kinematic variables of the reaching hand:  
11 cumulative trajectories (Figure 3B) and cumulative rotations (Figure 3C). The observation  
12 of the graphs suggests that the participants followed longer exploratory movements of the  
13 reaching right hand, with more cumulative rotation in the head mode than in the hand  
14 mode. Interestingly, the variability was more important in the head mode, and in particular  
15 regarding the cumulative hand rotations. ANOVAs were conducted with the factors Mode  
16 and Block on the dependent variables describing reaching hand trajectories: cumulative  
17 trajectories, number of velocity peaks, and cumulative rotations of the reaching hand  
18 trajectory. The results for cumulative trajectories failed to show any significant effect of  
19 Mode ( $F[1,11]=4.28$ ,  $p=0.062$ ) or Block ( $F[2,22]=1.12$ ,  $p=0.34$ ). Similarly, the results on  
20 number of velocity peaks failed to show any significant main effect of Mode  
21 ( $F[1,11]=0.45$ ,  $p=0.51$ ) or Block ( $F[2,22]=2.35$ ,  $p=0.11$ ). The results on cumulative  
22 rotations of the reaching hand trajectory showed a significant effect of Mode ( $F[1,11]=5.8$ ,  
23  $p=0.034$ ), no effect of Block ( $F[2,22]=0.6$ ,  $p=0.55$ ), and no interaction between Mode and  
24 Block ( $F[2,22]=0.71$ ,  $p=0.50$ ). There were more cumulative rotations of the reaching hand  
25 trajectory in the head mode than in the hand mode (see Figure 3C).

26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44 The direct distance followed by the head and its kinematic variables: cumulative  
45 trajectories and cumulative rotations are shown on Figure 3D-F. The direct distance  
46 followed by the head during reaching showed a significant effect of Mode ( $F[1,11]=82.06$ ,  
47  $p < 0.0001$ ), no effect of Block ( $F[2,22]=1.18$ ,  $p=0.32$ ), and no interaction between these  
48 two factors ( $F[2,22]=1.86$ ,  $p=0.17$ ). The distance was  $9.76 \pm 2.01$  cm in the hand mode  
49 and  $32.75 \pm 3.03$  cm in the head mode (see Figure 3D). The cumulative trajectories and  
50 rotations of the head were greater in the head mode ( $171.9 \pm 8.65$  cm and  $319.2 \pm 16.7^\circ$ ,  
51 respectively) than in the hand mode ( $44 \pm 2.74$  cm and  $74.5 \pm 3.38^\circ$ , respectively), with a  
52 higher variability in the head mode (see Figure 3E and 3F). The conditions for ANOVA  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 were not fully met (sphericity test  $p= 0.002$ ). Multivariate analysis was used with the  
5 factors Block and Mode on the cumulative trajectories. The analysis showed a significant  
6 effect of Mode ( $p<0.0001$ ), no effect of Block ( $p=0.1$ ), and no interaction between these  
7 two factors ( $p=0.06$ ). Two factors ANOVA for cumulative rotations (Figure 3F) showed a  
8 significant effect of Mode ( $F[1,11]=48.36$ ,  $p<0.0001$ ), no effect of Block ( $F[2,22]=0.98$ ,  
9  $p=0.39$ ), and no interaction between these two factors ( $F[2,22]=0.39$ ,  $p=0.68$ ).

### 16 17 *Learning and adaptation:*

18 As was described above, there was no effect of the session and the effect of Block  
19 appeared to be complex. Numerically, the success rate increased with Block (Figure 2A)  
20 but the statistics were difficult to interpret, in particular, a borderline sphericity test  
21 ( $p=0.044$ ) weakened the results of the ANOVA. Non parametric Friedman analysis  
22 showed an effect of Block in the head mode ( $p=0.011$ ) but not in the hand mode ( $p=0.30$ ).  
23 It is likely that an important inter-individual variability contributed to this complexity. A  
24 possible interpretation is that the lack of significant effect in the hand mode is due to a  
25 saturation of the success rate (8 out of 12 participants succeeded in at least 89% of the  
26 trials) inducing a ceiling effect, which did not occur in the head mode.

27  
28  
29  
30  
31  
32  
33  
34  
35 (Fig. 3 about here)

### 36 37 38 39 *Discussion of Experiment 1*

40 In Experiment 1, the participants were able to use auditory feedback in order to  
41 guide reaching movements. In addition, there was a significant improvement in success  
42 rate in the head mode and a ceiling effect in the hand mode. This result indicates that the  
43 participants were able to adapt to a new audio-motor environment and to learn efficient  
44 motor strategies. However, most of the kinematic variables did not vary with blocks. A  
45 possible explanation is that the participants used quite different individual strategies when  
46 completing motor tasks. For example, some of the participants performed relatively direct  
47 reaching movements, but most of them used various oscillatory exploratory movements  
48 that cannot be easily compared between them.

49 The second main result from Experiment 1 is that the success rate was higher in the  
50 hand mode than in the head mode. In addition, the kinematic analysis of the reaching  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 dominant hand showed that the target was reached after a simpler movement trajectory  
5 characterized by less cumulative rotations in the hand mode than in the head mode. This  
6 result indicates that when using a sensory substitution device, or any other technology, the  
7 idea that sensory organs would benefit from staying at their usual location is not  
8 necessarily true, given that here, the most unnatural condition was the most efficient.  
9

10  
11  
12  
13 It should however be underlined that there were two main differences between the  
14 hand mode and the head mode. First, in the hand mode, there was a spatial coincidence  
15 between the sensor and effector that was not there in the head mode. Second, the  
16 amplitude of possible movements was greater in the hand mode than in the head mode. As  
17 a consequence, a second experiment was conducted in order to investigate if these two  
18 differences account for the better performance in the hand mode than in the head mode.  
19  
20  
21  
22  
23  
24  
25

## 26 **Experiment 2**

27  
28 The better performance in the hand mode than in the head mode obtained in  
29 Experiment 1 can be due to the fact that the head condition involved more complexity and  
30 less amplitude of movements. First, in the hand condition, the sensor (the virtual listener)  
31 was located on the hand which was also used to catch the fly. Second, head movements  
32 have less amplitude than hand movements. To investigate the effect of these differences on  
33 participants' performance, in Experiment 2, in the hand condition, the listener was placed  
34 on the non-dominant hand while the dominant hand was used to reach the target. In  
35 addition, the participants' non-dominant hand displacements were constrained in order to  
36 have similar amplitude and degrees of freedom as that of the head.  
37  
38  
39  
40  
41  
42  
43  
44  
45

## 46 **Materials and methods**

### 47 *Participants*

48  
49  
50 18 new participants (9 females, age range: 20-50 years), recruited at the Paris  
51 Descartes University, completed Experiment 2. Inclusion criteria were being aged between  
52 18 and 50 years with no self-reported neurological disorder nor auditory impairment. This  
53 experiment was conducted in accordance with the ethical standards laid down in the 1991  
54 Declaration of Helsinki.  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 *Experimental set-up and procedure*  
5

6 The same devices and methods as in Experiment 1 were used in Experiment 2, but  
7 this time three sensors were used and placed on the two hands and the head, respectively.  
8 The reaching movements were again performed by the dominant hand. The listener was  
9 defined as the head sensor (head mode, as in Experiment 1) or as the non-dominant hand  
10 sensor (non-dominant hand mode). As in Experiment 1, the perceived sounds varied as a  
11 function of the sensor's position and orientation with respect to the sound source. In the  
12 non-dominant hand mode, the participants were instructed to lean their elbow on a  
13 homemade backing fixed on the armrest of the chair, with their forearm elevated so that  
14 the hand was at the shoulder's height. This posture aims at limiting the non-dominant  
15 hand's degrees of freedom, so that they are similar to those of the head. The conditions of  
16 presentation of each auditory target was the same as in Experiment 1: 9 different target  
17 positions were presented in each block, in a randomized order. There were 3 successive  
18 blocks for each of the two modes with the participants completing 54 trials in total.  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30

31 *Performance and motion analyses*  
32

33 The same dependent variables as in Experiment 1 were computed: success rate, trial  
34 duration, kinematics of head and both hands.  
35  
36  
37  
38

39 **Results**  
40

41 *Success rate*  
42

43 The success rate ranged between 14.8 and 77.8% in the hand mode and between  
44 37.0 and 88.9% in the head mode (mean  $\pm$  sem  $55.93 \pm 4.11$  %, in the hand mode;  $55.35 \pm$   
45  $3.54$  % in the head mode). As in Experiment 1, there was no effect of Order of sessions  
46 (Pearson's Chi-squared test,  $p=0.60$ ) and the results were pooled across sessions.  
47  
48  
49

50 Figure 4A illustrates the evolution of Success rate in the two modes. An ANOVA  
51 with the factors Mode and Block showed a significant of Block ( $F[2,32]=15.7$ ,  $p<0.0001$ ),  
52 no effect of Mode ( $F[1,16]=1.31$ ,  $p=0.27$ ), and no interaction between these two factors  
53 ( $F[2,32]=1.04$ ,  $p=0.36$ ). A Bonferoni Post-Hoc test showed that the success rates in Blocks  
54 2 and 3 were significantly higher than in Block 1 ( $p<0.001$ ).  
55  
56  
57  
58

59 (Fig. 4 about here)  
60  
61  
62  
63  
64  
65

### *Trial duration*

For Trial duration, the conditions for the ANOVA were not met and a Multivariate analysis failed to show any significant effect of Mode and Block ( $p > 0.5$ ), nor interaction between these two factors ( $p = 0.19$ ). The numerical trend (illustrated in Figure 4B) indicated a slight decrease in trial duration with Block, with an increased variability.

### *Kinematics of the reaching right hand*

Figure 5A-C illustrates the kinematic variables of the reaching dominant hand during the task. As in Experiment 1, the direct distance followed by the reaching hand did not vary with the factors Mode and Block, confirming that movement execution during successful trials was reproducible across conditions (see Figure 5A). The cumulative trajectories and cumulative rotations are shown in Figure 5 B,C. An ANOVA conducted with the factors Block and Mode on the cumulative trajectories indicated a borderline effect of Block ( $F[2,32] = 3.23$ ,  $p = 0.052$ ), no significant effect of Mode ( $F[1,16] = 3.27$ ,  $p = 0.08$ ), and no interaction between these two factors ( $F[2,32] = 1.0$ ,  $p = 0.037$ ). An ANOVA conducted with the factors Block and Mode on the cumulative rotations of the reaching hand did not reveal any significant effect of Mode ( $F[1,16] = 0.43$ ,  $p = 0.8$ ), nor Block ( $F[2,32] = 1.6$ ,  $p = 0.21$ ), nor interaction ( $F[2,32] = 0.18$ ,  $p = 0.83$ ). However, the ANOVA conducted on the number of velocity peaks showed a significant effect of Mode ( $F[1,16] = 8.36$ ,  $p = 0.01$ ), without any effect of Block ( $F[2,32] = 1.49$ ,  $p = 0.23$ ), nor interaction between these two factors ( $F[2,32] = 0.46$ ,  $p = 0.63$ ). The number of velocity peaks was higher in the head mode ( $18.2 \pm 1.76$ ) than in the hand mode ( $13.8 \pm 1.44$ ).

### *Kinematics of the listener*

The listener was the head or the non-dominant hand as a function of the condition. The kinematics of the listener are shown in Figure 5D-F. A two factors ANOVA on the direct distance followed by the listener (Figure 5D) showed a significant effect of Mode ( $F[1,16] = 8.48$ ,  $p = 0.01$ ), no effect of Block ( $F[2,32] = 0.14$ ,  $p = 0.63$ ), and no interaction between these two factors ( $F[2,32] = 0.611$ ,  $p = 0.54$ ). The non-dominant hand (blue line) moved  $14.6 \pm 1.37$  cm and the head (red line) moved  $21.7 \pm 1.91$  cm.

1  
2  
3  
4 An ANOVA with the factors Mode and Block showed that the listener's  
5 cumulative trajectories (Figure 5E) did not vary with Mode ( $F[1,16]=0.11$ ,  $p=0.74$ ), nor  
6 with Block ( $F[2,32]=0.24$ ,  $p=0.79$ ), and there was no interaction ( $F[2,32]=0.95$ ,  $p=0.39$ ). A  
7 two factors ANOVA showed that the cumulative rotations of the listener did not vary with  
8 Mode ( $F[1,16]=1.55$ ,  $p=0.23$ ) nor with Block ( $F[2,32]=0.46$ ,  $p=0.62$ ), however, there was  
9 a significant interaction between Block and Mode ( $F[2,32]=3.93$ ,  $p=0.03$ ). This suggests  
10 that the cumulative rotations decrease with Block only when the listener is fixed on the  
11 head (Figure 5F). This interaction was confirmed by Multivariate analysis ( $p=0.019$ ).  
12  
13  
14  
15  
16  
17  
18  
19  
20

### 21 *Discussion of Experiment 2*

22 Experiment 2 revealed that the participants were able to adapt their motor  
23 behaviour to a novel audio-motor environment that involves a new sensorimotor  
24 coordination between the two hands. In this experiment, there was no significant  
25 difference between the head mode and the hand mode, neither for the success rate nor for  
26 trial duration. These results suggest that the higher performance obtained in the hand mode  
27 in Experiment 1 was due to the spatial coincidence between the effector and the sensor.  
28 Surprisingly, the head mode that was supposed to be very close to a natural condition (for  
29 the human species, the ears are on the head) is not significantly better than the hand mode,  
30 and improvements are slower. As in Experiment 1, even if performance seems to be  
31 similar in the two modes, the participants moved their virtual ears with smaller amplitudes  
32 in the hand mode.  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43

### 44 **General discussion**

45 The study reported here investigated adaptation to a novel audio-motor  
46 environment. An electromagnetic system recorded the position and orientation of the  
47 participants' head and hand and it was connected to a 3D system that provides auditory  
48 feedback on movements. The participants' task was to reach an auditory target. The  
49 target's sound was perceived by using virtual ears (the listener) located either on the  
50 participants' hand (hand mode) or on their head (head mode). In Experiment 1, the  
51 participants' hand position was unconstrained and the listener was placed on the same  
52 hand as the one used to reach the target. To control for the role of spatial coincidence and  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 movement amplitude, in Experiment 2, hand movements were constrained in order to have  
5 similar amplitude and degrees of freedom as those of the head. In addition, the listener was  
6 placed on the other hand than the one used to reach the target.  
7  
8

9  
10 The first main result to emerge from our study was that the participants were able  
11 to localize a source within the 3D auditory environment. This result indicates that people  
12 are able to learn new efficient motor strategies and to adapt to a new audio-motor  
13 environment (see also Safstrom and Edin, 2006). The second main result to emerge from  
14 our study was the higher performance obtained when the sensor was located on the hand  
15 used to reach the target. This was the case in Experiment 1, in which performance was  
16 better in the hand mode than in the head mode. It should be underlined here that in the  
17 head mode, the task is mainly a binaural 3D sound source localization task; the reaching  
18 component of the task is secondary. The binaural localization of a virtual sound source is  
19 critically dependent on the algorithm used for spatialisation. In the present study it was not  
20 optimal since it was not personalized by Head Related Transfer Function (HRTF, see  
21 Simon *et al.*, 2016). On the other hand, auditory feedback in the hand mode is similar to  
22 movement sonification acting as a guide for the reaching hand toward the target  
23 (Bevilacqua *et al.*, 2016). In this hand mode, the task resembles more a guided approach  
24 task than a pure reaching task. In Experiment 2, the sensor was placed on the other hand  
25 and its movements were constrained to be similar to those of the head. The head mode was  
26 supposed to be closer to a natural condition, approximately the position of ears. However,  
27 performance was similar in the hand mode and in the head mode and improvements were  
28 slower in the head mode.  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43

44 This result contradicts some hypotheses from sensorimotor theories of perception  
45 according to which the use of a head-mounted camera shares more sensorimotor  
46 resemblance with vision than the use of a hand-held camera and thereby might favour  
47 better performance (e.g., O'Regan, 2011; O'Regan and Noë, 2001). However, it might be  
48 the case that sensorimotor resemblance influences experience on a phenomenological  
49 rather than at the behavioural level. It would be interesting to further investigate those  
50 factors differentially influencing the behavioural and phenomenological components of the  
51 task, as well as the individual differences potentially at play. Indeed, the extent to which  
52 the use of a sensory substitution device induces supplemental visual, auditory, tactile, or  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 cognitive processing varies from one user to another (e.g., Auvray et al., 2007-a; see also  
5  
6 Arnold *et al.*, 2017, Deroy and Auvray, 2014, for discussions on this topic). Thus, in our  
7  
8 study, performance appeared to be less linked to the natural aspect related to sensor  
9  
10 location than to the spatial coincidence between the body part performing the motor task  
11  
12 and the origin of the audio-motor coupling. These results have practical implications,  
13  
14 suggesting that under degraded perceptual situations (real or virtual), it could be more  
15  
16 efficient to use spatially coincident sensors and effectors.

17 The third main result from our study concerns the kinematic data, which allow  
18  
19 investigating the quality of movement associated with task success. Overall, the  
20  
21 kinematics of the reaching hand is not more informative than the success rate. However,  
22  
23 the kinematics of the head and non-dominant hand allows a better understanding of the  
24  
25 listener's behaviours of active exploration to perform the perceptual task. More precisely,  
26  
27 in Experiment 1, the analysis of head movements confirmed that the reaching task was  
28  
29 executed differently in the two modes. In the hand mode, the head moved around 10 cm  
30  
31 during reaching with relatively small cumulative trajectories and rotations. This  
32  
33 displacement can be at least partially explained by the known contribution of the upper  
34  
35 trunk to reaching movements (Mark *et al.*, 1997). In contrast, in the head mode, the head  
36  
37 roamed the workspace with much larger cumulative trajectories and rotations. The direct  
38  
39 3D displacement of the head at the end of reaching attained 30 cm which is a strikingly  
40  
41 large distance. This observation suggests that in the head mode, the function of the head is  
42  
43 not that of a stable egocentric reference frame for planning and executing reaching  
44  
45 movement (see also Crawford *et al.*, 2004; Paillard, 1991). Rather, it suggests that head  
46  
47 movements, guided by the auditory feedback, are primarily used to explore the 3D  
48  
49 workspace in order to ease target detection. We hypothesize that head movements simulate  
50  
51 hand movements toward the target. Then, the position and orientation of the head toward  
52  
53 the target (heading) can be secondarily used to perform target-directed hand reaching  
54  
55 movements (see also Boyer *et al.*, 2013).

56 Although the two experiments involved differences in terms of constraints on  
57  
58 movements and location of the listener, the kinematics of the listener had similarities in  
59  
60 the head during head mode in Experiment 1, the head, and the non-dominant hand in  
61  
62 Experiment 2, which have similar constraints in terms of amplitude. Indeed, in these three  
63  
64  
65

1  
2  
3  
4 situations, there were large cumulative trajectories and rotations, contrasting with the  
5 relatively limited movements of the right reaching hand in hand mode in Experiment 1. In  
6 Experiment 2, the kinematics of the head and non-dominant hand are similar. However,  
7 the direct distance followed during reaching is smaller when the listener is fixed on the  
8 non-dominant hand (around 15 cm) than when it is fixed on the head (around 22 cm). This  
9 is quite paradoxical since the movement of the non-dominant hand is free (elbow and wrist  
10 degrees of freedom) and probably less effortful than the trunk and neck movements  
11 needed to move the head listener. These results suggest that the role of the head to explore  
12 space and to indicate target direction (heading) does not transfer easily to the hand.  
13  
14  
15  
16  
17  
18  
19  
20

21 Participants' performance suggests that the paradigm used in our study can provide  
22 useful 3D auditory feedback. This can be applied to a broad range of situations. In the field  
23 of sensory substitution, our results are in line with many other studies investigating the  
24 relevance of using audition to convey visual information (e.g., the vOICe, Meijer, 1992;  
25 Vibe, Hanneton et al., 2010; the EyeMusic, Levy-Tzedek et al., 2014; the See ColOr;  
26 Bologna et al., 2009). They additionally suggest that we can step away from reproducing  
27 the conditions of natural perception and rather take into account how motor performance  
28 relates the best to changes in sensory information. A similar audio-motor coupling could  
29 also be used to compensate proprioceptive deficits, for which tactile devices have been  
30 mainly used so far (e.g., see Danilov et al., 2007; Diot et al., 2014). In addition, augmented  
31 auditory feedback in a game-like situation may be a useful tool for rehabilitation of  
32 sensory-motor functions (Schmitz *et al.*, 2018). The hand mode could thus be used as a  
33 way of providing augmented feedback to patients with movement disorders such as  
34 hemiparesis following stroke. Indeed, due to the strong link between sensory inputs and  
35 motor outputs, augmented auditory feedback can be a useful tool in rehabilitation. This  
36 paradigm can also be used to train specific technical skills such as minimally invasive  
37 surgery which relies now on virtual reality simulators with enhanced sensory feedback  
38 (see Papanikolaou *et al.*, 2019, for a review). Finally, audio-motor feedback can be useful  
39 in the field of perceptual augmentation, in which devices are developed to improve access  
40 to visual information in a variety of situations, for instance when there is an excess of  
41 visual information to process, like when driving a car (e.g., Ho and Spence, 2008) or when  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 conditions of perception are degraded, as can be the case for those involved in firefighting  
5  
6 (Carton and Dunne, 2013) or military operations (Jones et al., 2009).  
7  
8  
9

## 10 **Acknowledgements**

11 We thank Ross Parry for his reading and useful comments on the manuscript. This work  
12 was supported by the Mission pour l'Interdisciplinarité (CNRS, Auton, Sublima Grant)  
13 and Labex SMART (ANR-11-IDEX-0004-02) supported by French state funds managed  
14 by the ANR within the Investissements d'Avenir program under reference ANR-11-IDEX-  
15 0004-02.  
16  
17  
18  
19  
20  
21  
22

## 23 **Footnote**

24  
25 1. See IA-SIG Interactive 3D Audio Rendering and Evaluation Guidelines (Level 1) and  
26 OpenAL Documentation for details: <https://www.openal.org/>  
27  
28  
29  
30

## 31 **References**

- 32 Ahissar, M. and Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual  
33 learning. *Trends Cogn. Sci.* **8**, 457–464.  
34  
35 Arno, P., Vanlierde, A., Streel, E., Wanet Defalque, M. C., Sanabria-Bohorquez, S. and  
36 Veraart, C. (2001). Auditory substitution of vision: pattern recognition by the blind.  
37 *Applied Cogn. Psychol.* **15**, 509–519.  
38  
39 Arnold, G. and Auvray, M. (2018). Tactile recognition of visual stimuli: Specificity versus  
40 generalization of perceptual learning. *Vision Res.* **152**, 40–50.  
41  
42 Arnold, G. and Auvray, M. (2014). Perceptual learning: Tactile letter recognition transfers  
43 across body surfaces. *Multisens. Res.* **27**, 71–90.  
44  
45 Arnold, G., Pesnot-Lerousseau, J. and Auvray, M. (2017). Individual differences in sensory  
46 substitution. *Multisens. Res.* **6**, 579–600.  
47  
48 Auvray, M., Hanneton, S. and O'Regan, J. K. (2007-a). Learning to perceive with a visuo-  
49 auditory substitution system: Localization and object recognition with The Voice.  
50 *Perception* **36**, 416–430.  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

- 1  
2  
3  
4 Auvray, M. and Myin, E. (2009). Perception with compensatory devices. From sensory  
5 substitution to sensorimotor extension. *Cogn. Sci.* **33**, 1036–1058.  
6  
7  
8 Auvray, M., Philipona, D., O'Regan, J. K. and Spence, C. (2007-b). The perception of  
9 space and form recognition in a simulated environment: The case of minimalist sensory-  
10 substitution devices. *Perception* **36**, 1736–1751.  
11  
12  
13  
14 Bach-y-Rita, P. and Kercel, S. W. (2003). Sensory substitution and the human-machine  
15 interface. *Trends Cogn. Sci.* **7**, 541–546.  
16  
17  
18 Bach-y-Rita, P., Collins, C. C., Saunders, F. A., White, B. and Scadden, L. (1969). Vision  
19 substitution by tactile image projection. *Nature* **221**, 963–964.  
20  
21  
22  
23 Bevilacqua, F., Boyer, E. O., Françoise, J., Houix, O., Susini, P., Roby-Brami, A. and  
24 Hanneton S. (2016). Sensori-motor learning with movement sonification: Perspectives  
25 from recent interdisciplinary studies. *Front. Neurosci.* **10**:385.  
26  
27  
28  
29 Bernstein, N. (1967). *The coordination and regulation of movements*. Oxford: Pergamon  
30 Press.  
31  
32  
33 Bologna, G., Deville, B. and Pun, T. (2009). On the use of the auditory pathway to  
34 represent image scenes in real-time, *Neurocomputing* **72**, 839–849.  
35  
36  
37 Boyer, E.O., Babayan, B. M., Bevilacqua, F., Noisternig, M., Warusfel, O., Roby-Brami,  
38 A., Hanneton, S. and Viaud-Delmon, I. (2013). From ear to hand: the role of the  
39 auditory-motor loop in pointing to an auditory source. *Front. Comput. Neurosc.* **7**: 26.  
40  
41  
42  
43 Brown, D. J., Macpherson, T. and Ward, J. (2011). Seeing with sound? Exploring different  
44 characteristics of a visual-to-auditory sensory substitution device. *Perception* **40**, 1120–  
45 1135.  
46  
47  
48  
49 Carton, A. and Dunne, L. E. (2013). Tactile distance feedback for firefighters: design and  
50 preliminary evaluation of a sensory augmentation glove. In *Proceedings of the 4th*  
51 *Augmented Human International Conference (AH '13)*. ACM, New York, NY, USA, 58-  
52 64.  
53  
54  
55  
56  
57 Chebat, D. R., Maidenbaum, S. and Amedi A. (2015). Navigation using sensory  
58 substitution in real and virtual mazes. *PLoS ONE* **10**(6), e0126307.  
59  
60  
61  
62  
63  
64  
65

- 1  
2  
3  
4 Chebat, D. R., Schneider, F. C., Kupers, R. and Ptito, M. (2011). Navigation with a sensory  
5 substitution device in congenitally blind individuals. *Neuroreport* **22**, 342–347.  
6  
7  
8 Crawford, J. D., Medendorp, W. P. and Marotta, J. J (2004). Spatial transformations for  
9 eye-hand coordination. *J Neurophysiol.* **92**, 10–19.  
10  
11  
12 Danilov, Y., Tyler, M., Skinner, K., Hogle, R. and Bach-y-Rita, P. (2007). Efficacy of  
13 electrotactile vestibular substitution in patients with peripheral and central vestibular  
14 loss, *J. Vestib. Res.* **17**, 119–130.  
15  
16  
17  
18 Deroy, O. and Auvray, M. (2014). A crossmodal perspective on sensory substitution, in  
19 *Perception and Its Modalities*, S. Biggs, M. Matthen, & D. Stokes (Eds.), pp. 327–349,  
20 Oxford: Oxford University Press.  
21  
22  
23  
24 Diot, B., Halavackova, P., Demongeot, J. and Vuillerme, N. (2014). Sensory substitution  
25 for balance control using a vestibular-to-tactile device, *Multisens. Res.* **27**, 313–336.  
26  
27  
28  
29 Gallay, M., Denis, M. and Auvray, M. (2013). Navigation assistance for blind pedestrians:  
30 Guidelines for the design of devices and implications for spatial cognition, in  
31 *Representing space in cognition: Interrelations of behaviour, language, and formal*  
32 *models*, T. Tenbrink, J. Wiener, & C. Claramunt (Eds.), pp. 244–267, Oxford: Oxford  
33 University Press.  
34  
35  
36  
37  
38  
39 Haigh, A., Brown, D. J., Meijer, P. and Proulx, M. J. (2013). How well do you see what  
40 you hear? The acuity of visual-to-auditory sensory substitution. *Fron. Psychol.* **4**, 330.  
41  
42  
43 Hanneton, S., Auvray, M. and Durette, B. (2010). The Vibe: A versatile vision-to-audition  
44 sensory substitution device. *Appl Bionics Biomech.* **7**, 269–276.  
45  
46  
47 Hanneton, S., Herquel, P. and Auvray, M. (2015). Intermodal recoding of a video game:  
48 Learning to process signals for motion perception in a pure auditory environment. *Int J*  
49 *Adapt Control Signal Process* **29**, 1475–1483.  
50  
51  
52  
53 Ho, C. and Spence, C. (2008). *The Multisensory Driver: Implications for Ergonomic Car*  
54 *Interface Design*. Ashgate Publishing, Aldershot, UK.  
55  
56  
57 Kim, J. K. and Zatorre, R. J. (2008). Generalized learning of visual-to-auditory substitution  
58 in sighted individuals. *Brain Res.* **1242**, 263–275.  
59  
60  
61  
62  
63  
64  
65

- 1  
2  
3  
4 Kim, J. K. and Zatorre, R. J. (2010). Can you hear shapes you touch? *Exp. Brain Res.* **202**,  
5 747–754.  
6  
7  
8 Levy-Tzedek, S., Hanassy, S., Abboud, S., Maidenbaum, S. and Amedi, A. (2012). Fast,  
9 accurate reaching movements with a visual-to-auditory sensory substitution device.  
10 *Restor Neurol Neurosci.* **30**, 313–323.  
11  
12  
13 Levy-Tzedek, S., Riemer, D. and Amedi, A. (2014). Color improves “visual” acuity via  
14 sound, *Front. Neurosci.* **8**, 358.  
15  
16  
17 Marks, D. F. (1973). Visual imagery in the recall of pictures. *Br J Psychol.* **64**, 17–24.  
18  
19  
20 Mark, L. S., Nemeth, K., Gardner, D., Dainoff, M. J., Paasche, J., Duffy, M. and Grandt, K.  
21 (1997). Postural dynamics and the preferred critical boundary for visually guided  
22 reaching. *J Exp Psychol Hum Percept Perform.* **23**, 1365–1379.  
23  
24  
25 Meijer, P. B. L. (1992). An experimental system for auditory image representations. *IEEE*  
26 *T Bio-Med Eng* **39**, 112–121.  
27  
28  
29 Novak, K. E., Miller, L. E. and Houk, J. C. (2003). Features of motor performance that  
30 drive adaptation in rapid hand movements. *Exp. Brain Res.* **148**, 388–400.  
31  
32  
33 O'Regan, J. K. (2011). *Why Red Doesn't Sound Like a Bell. Understanding the Feel of*  
34 *Consciousness*. Oxford: Oxford University Press.  
35  
36  
37 O'Regan, J. K. and Noë, A. (2001). A sensorimotor account of vision and visual  
38 consciousness. *Behav Brain Sci.* **24**, 939–973.  
39  
40  
41 Papanikolaou, I.G., Haidopoulos, D., Paschopoulos, M., Chatzipapas, I., Loutradis, D. and  
42 Vlahos, N.F. (2019). Changing the way we train surgeons in the 21th century: A  
43 narrative comparative review focused on box trainers and virtual reality simulators, *Eur*  
44 *J Obstet Gynecol Reprod Biol.* **235**, 13–18.  
45  
46  
47 Paillard, J. (1991). Motor and representational framing of space. In J. Paillard (Ed.), *Brain*  
48 *and Space* (pp. 163-182). New York, US: Oxford University Press.  
49  
50  
51 Pollok, B., Schnitzler, I., Mierdorf, T., Stoerig, P. and Schnitzler, A. (2005). Image-to-  
52 sound conversion: Experience-induced plasticity in auditory cortex of blindfolded  
53 adults. *Exp. Brain Res.* **167**, 287–291.  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

- 1  
2  
3  
4 Proulx, M.J., Brown, D. J., Pasqualotto, A and Meijer, P. (2014). Multisensory perceptual  
5 learning and sensory substitution. *Neurosci Biobehav R* **41**, 16–25.  
6  
7  
8 Proulx, M. J., Stoerig, P., Ludowig, E. and Knoll, I. (2008). Seeing ‘where’ through the  
9 ears: effects of learning-by-doing and long-term sensory deprivation on localization  
10 based on image-to-sound substitution. *PloS one* **3**, e1840.  
11  
12  
13  
14 Ptito, M., Iversen, K., Auvray, M., Deroy, O. and Kupers, R. (2018). Limits of the classical  
15 functionalist perspective on sensory substitution, in *Sensory Substitution and*  
16 *Augmentation*, F. Macpherson (Ed.), London: Proceedings of the British Academy.  
17  
18  
19  
20 Renier, L., Laloyaux, C., Collignon, O., Tranduy, D., Vanlierde, A., Bruyer, R. and De  
21 Volder, A. G. (2005). The Ponzo illusion using auditory substitution of vision in sighted  
22 and early blind subjects. *Perception* **34**, 857–867.  
23  
24  
25  
26 Rand, M. K., Shimansky, Y., Stelmach, G. E. and Bloedel, J. R. (2004). Adaptation of  
27 reach-to-grasp movement in response to force perturbations. *Exp. Brain Res.* **154**, 50–  
28 65.  
29  
30  
31  
32 Safstrom, D. and Edin, B. B. (2006). Acquiring and adapting a novel audiomotor map in  
33 human grasping. *Exp. Brain Res.* **173**, 487–497.  
34  
35  
36  
37 Sampaio, E., Maris, S. and Bach-y-Rita, P. (2001). Brain plasticity: ‘Visual’ acuity of blind  
38 persons via the tongue. *Brain Res.* **908**, 204–207.  
39  
40  
41 Schmitz, G., Bergmann, J., Effenberg, E.O., Krewer, C., Hwang, T.H. and Müller, F.  
42 (2018). Movement sonification in stroke rehabilitation. *Fron. Neurol.* **9**: 389.  
43  
44  
45 Simon, L.S., Zacharov, N. and Katz, B. F. (2016). Perceptual attributes for the comparison  
46 of head-related transfer functions. *J. Acoust. Soc. Am* **140**, 3623–3632.  
47  
48  
49 Striem-Amit, E. and Amedi, A. (2014). Visual cortex extrastriate body-selective area  
50 activation in congenitally blind people “seeing” by using sounds. *Curr. Biol* **6**, 687–  
51 692.  
52  
53  
54  
55 Striem-Amit, E., Guendelman, M. and Amedi, A. (2012). ‘Visual’ acuity of the  
56 congenitally blind using visual-to-auditory sensory substitution. *PloS one* **7**, e33136.  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

Vandenberg, S. G. (1971). *Mental Rotation Test*. University of Colorado: Boulder, Colorado.

1  
2  
3  
4 **Figure captions**

5  
6 **Figure 1.**

7 Schematic representation of the experimental set-up. The auditory target is represented by  
8 a fly. For each mode, the sensor (or listener) is represented by a black square placed either  
9 on the dominant hand (A) or on the head (B). The schema illustrates how the participants'  
10 hand or head orientation relative to the sound source is used to produce spatialised  
11 auditory feedback.  
12  
13  
14  
15  
16  
17

18  
19 **Figure 2.**

20 Success rate and trial duration in Experiment 1. The graphs show the evolution of the  
21 success rate (A) and trial duration (B) during the three successive blocks for the two  
22 modes: hand mode (in blue) and head mode (in red). The symbols and vertical lines  
23 represent the mean and standard error to the mean.  
24  
25  
26  
27  
28

29  
30 **Figure 3.**

31 Kinematics of the reaching hand and of the head in the hand mode (in blue) and head  
32 mode (in red). The left part of the figure shows the kinematic variables of the reaching  
33 hand: 3D direct distance (A), cumulative trajectories (B), and cumulative rotation (C). The  
34 right part of the figure shows the kinematic variables of the head: 3D direct distance (D),  
35 cumulative trajectory (E), and cumulative rotation (F). The symbols and vertical lines  
36 represent the mean and standard error to the mean.  
37  
38  
39  
40  
41  
42  
43

44  
45 **Figure 4.**

46 Success rate and trial duration in Experiment 2. The graphs show the evolution of the  
47 success rate (A) and trial duration (B) during the three successive blocks for the two  
48 modes: hand mode (in blue) and head mode (in red). The symbols and vertical lines  
49 represent the mean and standard error to the mean.  
50  
51  
52  
53  
54

55  
56 **Figure 5.**

57 Kinematics of the reaching hand and of the head in the non-dominant hand mode (in blue)  
58 and in the head mode (in red). The left part of the figure shows the kinematic variables of  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

the reaching hand: 3D direct distance (A), cumulative trajectories (B), and cumulative rotations (C ). The right part of the figure shows the kinematic variables of the listener: 3D direct distance (D), cumulative trajectories (E), and cumulative rotations (F). The symbols and vertical lines represent the mean and standard error to the mean.

Figure 1

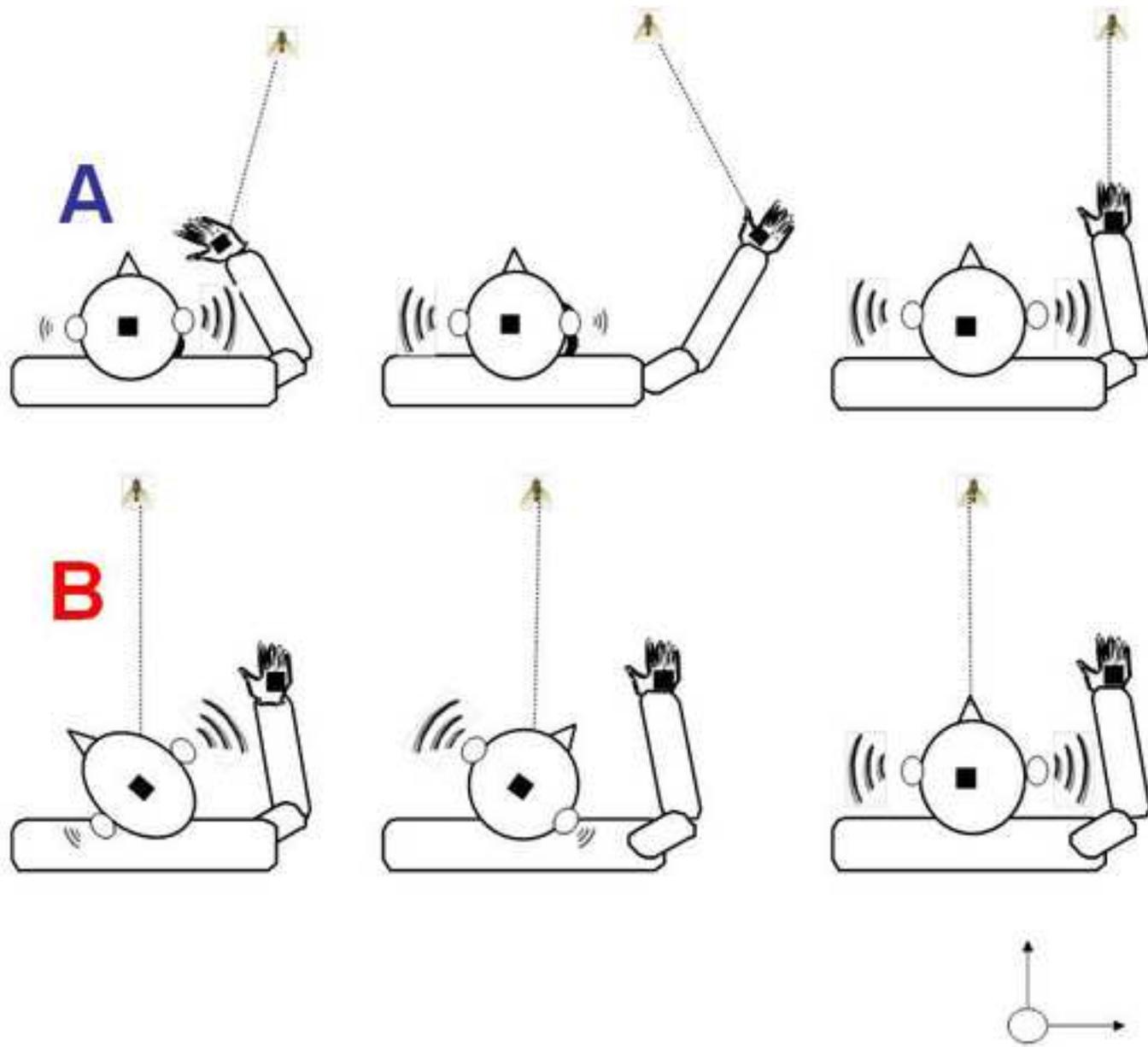


Figure 2

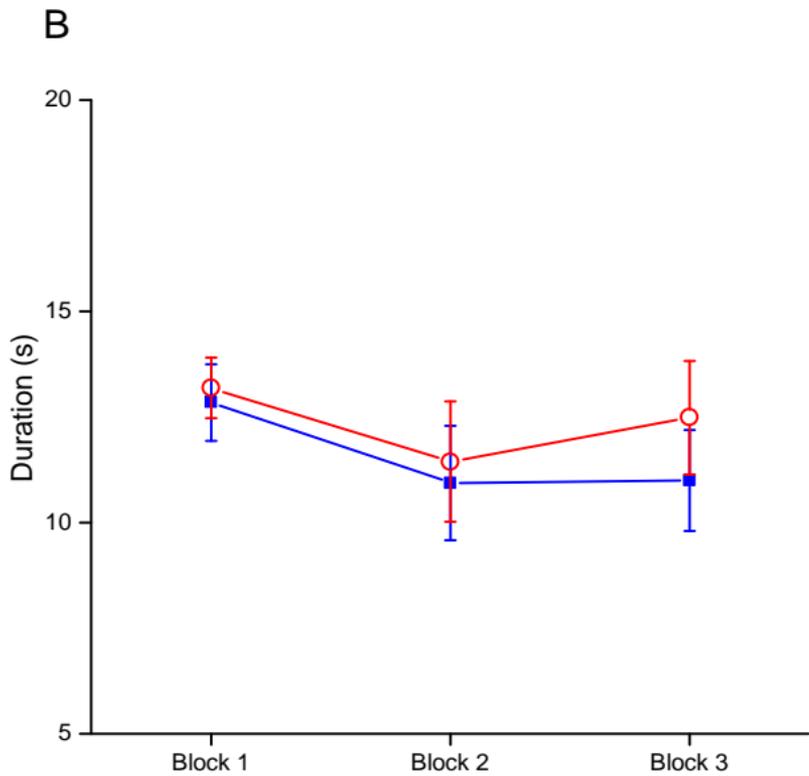
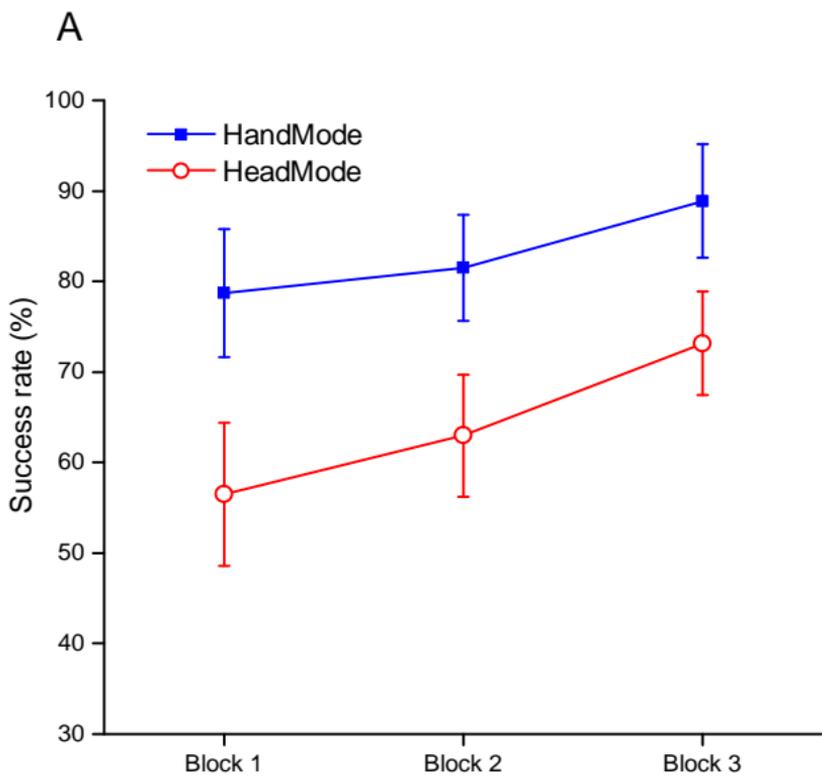


Figure 3

## Reaching hand

## Head

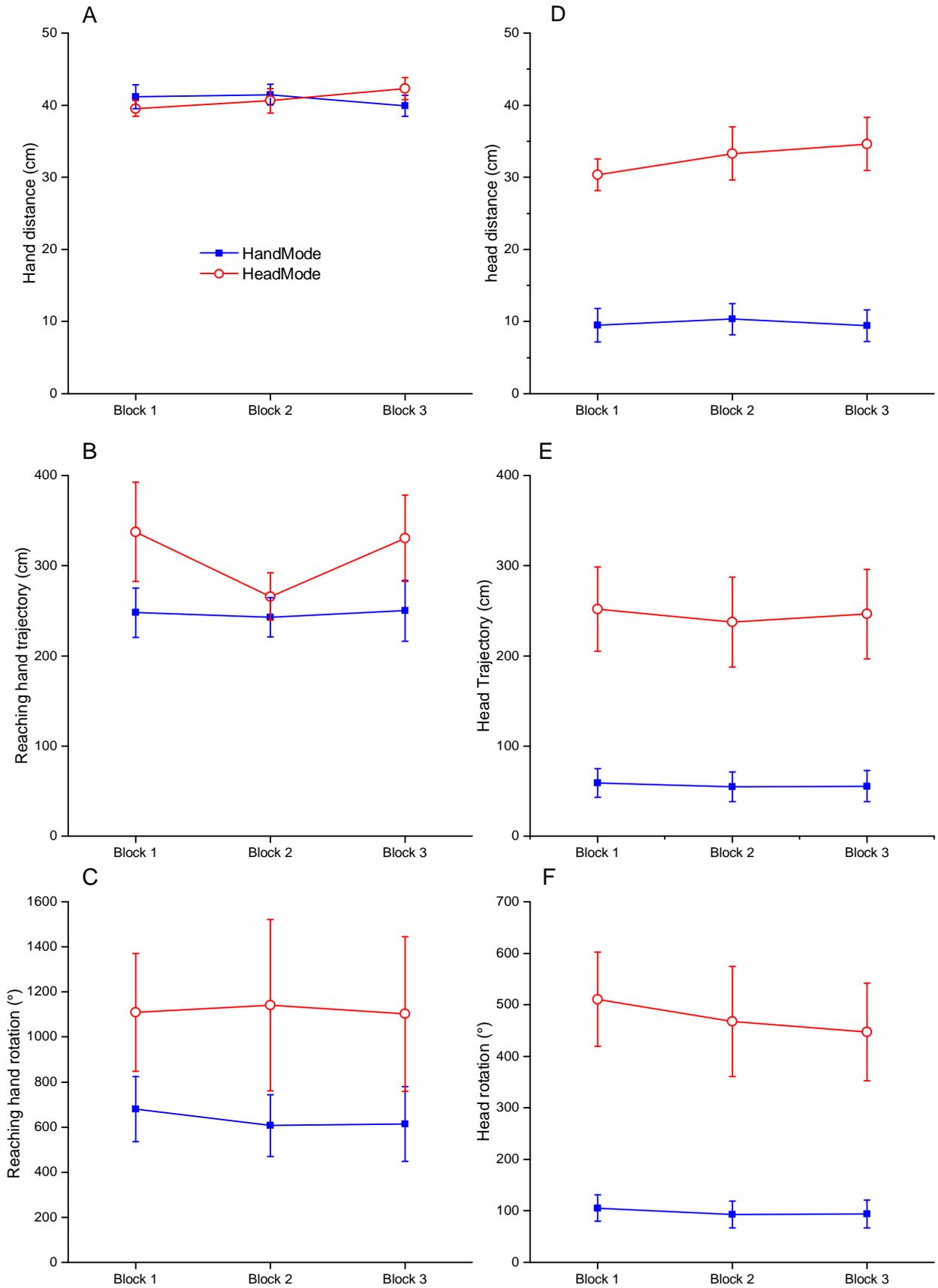


Figure 4

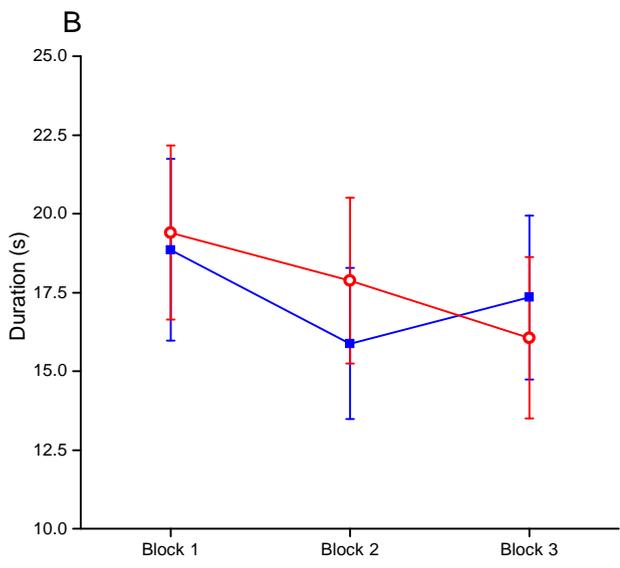
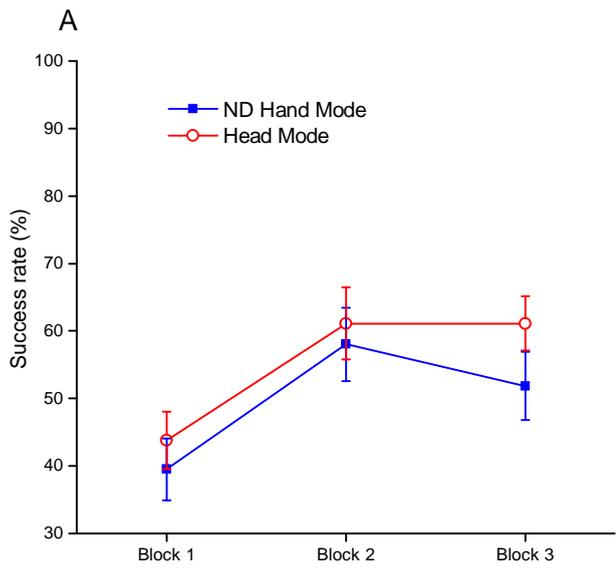
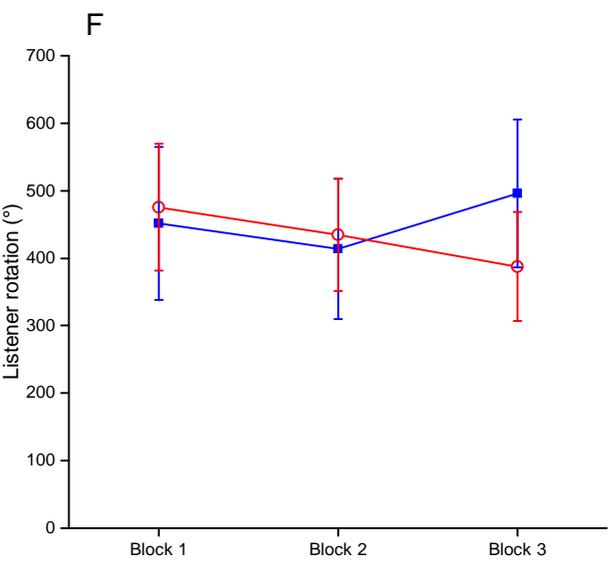
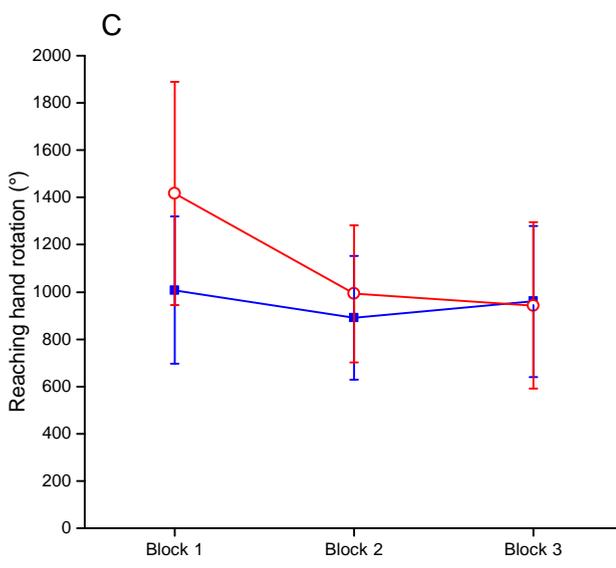
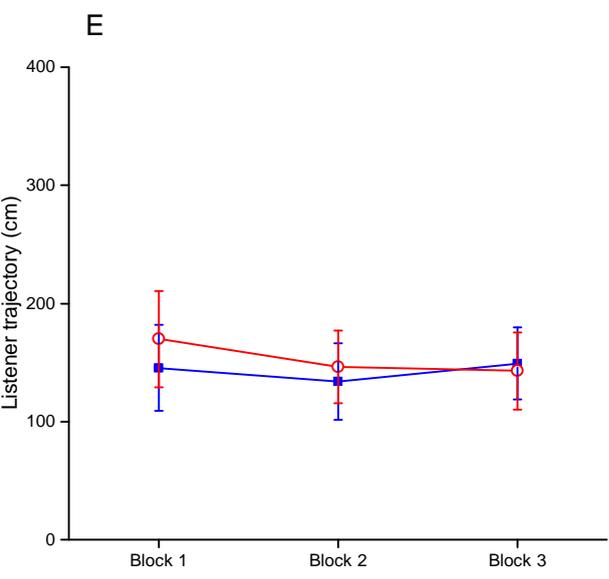
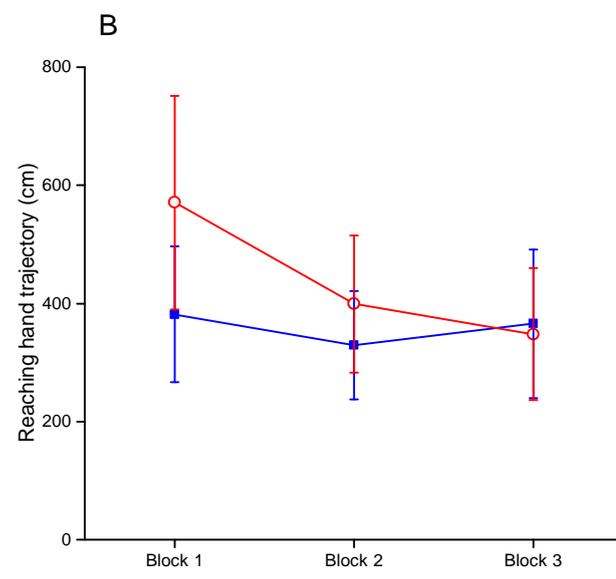
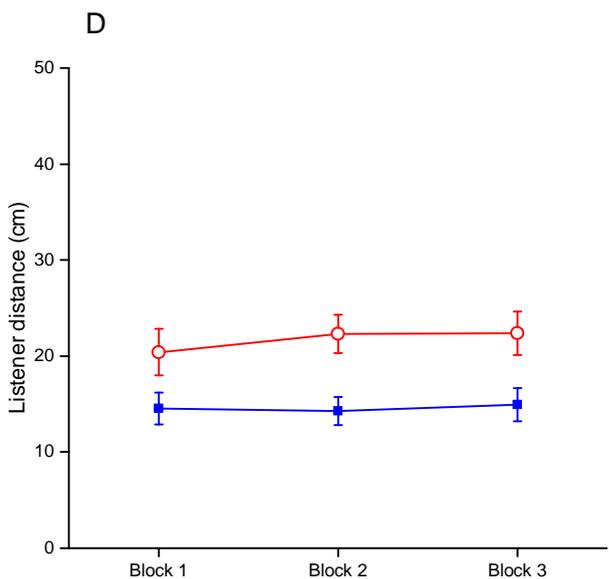
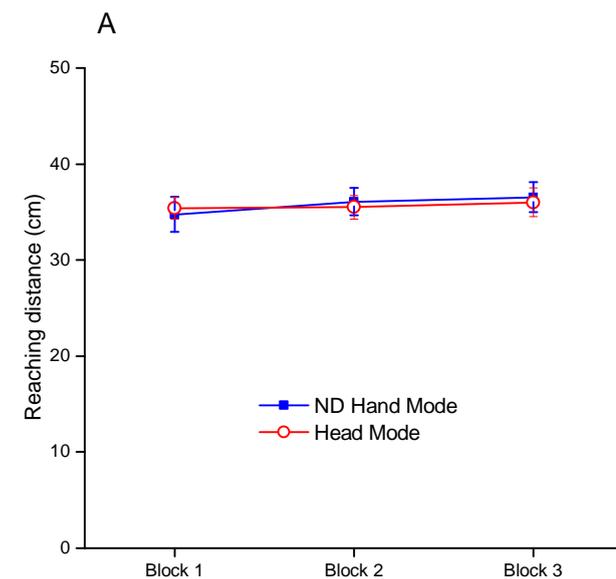


Figure 5

# Reaching hand

# Listener



**E**