



**HAL**  
open science

# Multisensory spatial perception and sensory substitution

Malika Auvray

► **To cite this version:**

Malika Auvray. Multisensory spatial perception and sensory substitution. Cognitive science. Sorbonne Université, 2018. tel-03975339

**HAL Id: tel-03975339**

**<https://hal.science/tel-03975339>**

Submitted on 6 Feb 2023

**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.



Distributed under a Creative Commons Attribution 4.0 International License



Sorbonne Université

Institut des Systèmes Intelligents et de Robotique

# Habilitation à Diriger des Recherches

Specialty: Biology

by Malika Auvray

## **Multisensory spatial perception and sensory substitution**

Viva planned for the 25<sup>th</sup> September 2018, before the following committee:

Dr Michel Denis	LIMSI, Orsay, France	(Examiner)
Pr Vincent Hayward	Sorbonne Université, Paris, France	(Examiner)
Dr. Christophe Lopez	LNSC, Marseille, France	(Referee)
Dr Michael J. Proulx	University of Bath, UK	(Referee)
Pr Maurice Ptito	School of Optometry, Montreal, Canada	(Examiner)
Pr Charles Spence	Oxford University, UK	(Examiner)
Dr Isabelle Viaud-Delmon	IRCAM, Paris, France	(Referee)

## **Acknowledgements**

I first would like to thank Michel Denis, Vincent Hayward, Christophe Lopez, Michael Proulx, Maurice Ptito, Charles Spence, and Isabelle Viaud-Delmon for agreeing to be part of my HDR committee.

Most of the research presented in this manuscript was conducted by the students and post-doctoral researchers that I have supervised and I want to warmly thank them for their work and for all the things I had the chance to learn from them. I would like to address special thanks to Gabriel Arnold, who stayed under my supervision for 5 years, and to Jessica Hartcher O'Brien and Laura Fernandez who stayed two years. Many thanks as well to the two PhD students Jérémy Roque and Adrien Girard, and to all the master students.

The research I have conducted has always been the fruit of collaborations, and I want to thank my colleagues for the chance I had to work with them, all the work and enthusiastic discussions we had together. I would like to express my deep gratitude to my collaborators, Vincent Hayward, Charles Spence, Michel Denis, Kevin O'Regan, Frédérique de Vignemont, Jérôme Dokic, Pierre Jacob, Ophelia Deroy, Christophe Joffrais, Sylvain Hannequin, Agnès Roby-Brami, Philippe Tarroux, Alberto Gallace, Barry Smith, Maurice Ptito, Ron Koppers, Charles Lenay, Erik Myin, Gaëtan Parseihian, Marieke Rohde, Mathieu Dubois. I also thank the researchers I more recently had the opportunity to work with, Fabrice Sarlegna, Nathanaël Jarasse, Wael Watcha, Alessandro Farnè, Mehdi Ammi, Jérémy Lafraire.

I would also like to thank many persons who provided me with invaluable advices and help along these years: Elisabeth Pacherie, Jean-Louis Vercher, Christian Lorenzi, Roberto Casati, Pascal Sommer, Raja Chatila, Sharon Peperkamp, Alva Noë, François Recanati, Joelle Proust, Vincent Gaudefroy.

I would like to thank the heads of the institutions I've successively been in for welcoming me and allowing me to pursue these lines of research. Finally, I would like to thank my colleagues from these institutions: Laboratoire Psychologie de la perception, Université de Technologie de Compiègne, Department of Experimental Psychology in Oxford, Department of Philosophy in Antwerp, Laboratory Neurophysics and Physiology, Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur, Institut Jean Nicod, and Institut des Systèmes Intelligents et de Robotique, Paris for sharing lab life with me and bringing to it joy and enriching discussions.

## Table of contents

I. Presentation of the scientific work.....	1
I.1. Abstract in English.....	1
I.2. Abstract in French.....	2
I.2. General introduction .....	3
II. Sensory substitution .....	4
II.1. Conceptual understanding of sensory substitution .....	5
II.2. Seeing the thunder while still hearing it .....	9
II.3. Sensory substitution and learning .....	13
II.4. Crossmodal correspondences for the optimization of the devices .....	17
II.5. Prototypes of sensory substitution devices .....	20
II.6. Conclusion on sensory substitution .....	21
III. Multisensory spatial reference frames .....	22
III.1. The variability in our reference frames .....	22
III.2. Adopting a reference frame, a natural process or a mere cognitive choice? .....	26
III.3. The influence of vision on the reference frames that are adopted.....	29
III.4. The embodied nature of spatial perspective taking .....	31
III.5. The role of proprioceptive information on spatial perspective taking .....	32
III.6. The transfer of a learned reference frame.....	33
III.7. The bi-stable palm .....	35
III.8. The processes of distal attribution .....	37
III.9. Conclusion on multisensory reference frames .....	40
IV. Perspectives for future research.....	42
IV.1. Exteriorisation and embodiment .....	42
IV.2. Sensory substitution of vision and proprioception .....	44
IV.3. Interactions between spatial and mental perspectives .....	45
V. References .....	46
CV .....	53
Annex. Selection of articles .....	75
1. Pesnot-Lerousseau, J., Arnold, G., & Auvray, M. (submitted). Visualizing sounds: training-induced plasticity with a visual-to-auditory conversion device.	
2. Arnold, G., & Auvray, M. (2014). Perceptual learning: Tactile letter recognition transfers across body surfaces. <i>Multisensory Research</i> , 27, 71-90.	
3. Arnold, G., & Auvray, M. (2017). Tactile recognition of visual stimuli: Specificity versus generalization of perceptual learning. <i>Vision Research</i> .	

4. Deroy, O., Fasiello, I., Hayward, V., & Auvray, M. (2016). Differentiated audio-tactile correspondences in sighted and blind individuals. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 1204-1214.
5. Arnold, G., Spence, C., & Auvray, M. (2016). Taking someone else's spatial perspective: Natural stance or effortful decentring? *Cognition*, *148*, 27-33.
6. O'Brien, J., & Auvray, M. (2016). Cognition overrides orientation dependence in tactile viewpoint selection. *Experimental Brain Research*, *234*, 1885-1892.
7. O'Brien, J., & Auvray, M. (2014). The process of distal attribution illuminated through studies of sensory substitution. *Multisensory Research*, *27*, 421-441.

# **I. Presentation of the scientific work**

## **I.1. Abstract**

My research focuses on perceptual processes and in particular on the relations between sensory modalities, both in their correspondences and in their interactions. The conducted research is based on an intrinsic multidisciplinary approach. The tested hypotheses come from a joint consideration of theoretical models, experimental and clinical approaches, with a preoccupation of applied outputs in the field of perceptual rehabilitation. The first axis of my research concerns sensory substitution devices, which aim at compensating one or several functions of a deficient sensory modality by means of information transmitted to an intact sensory modality. My research lies both upstream and downstream the design of sensory substitution devices. In particular, my work involves a conceptual understanding of the nature and possibilities offered by sensory substitution devices, studies of the central factors underlying their learning, their adequacy to the target population, taking into account their individual specificities, and the possibility of improving their design by the use of crossmodal correspondences. I also have collaborated to the design and tests of three sensory substitution devices. The second axis of my work concerns the multisensory groundings of our spatial representations, which vary as a function of the sensory modalities that are involved. Indeed, visual information is initially coded in retinotopic coordinates, auditory information is initially coded according to a head-centred reference frame, and tactile information is coded either in somatotopic or in head-centred coordinates. My aim is to understand how these reference frames interact to give rise to a functional representation of the environment and how these reference frames can be modified by the use of sensory substitution devices and other tools. Taken together these two research axes aim, at the fundamental level, at contributing to the comprehension of the plasticity of our perceptual and bodily representations during tool use. At the applied level, they aim at improving the design and learning of sensory substitution devices.

### **I.3. Résumé**

Mes recherches portent sur les processus perceptifs et en particulier sur la relation entre les modalités sensorielles dans leurs correspondances ainsi que dans leurs interactions. L'accent est mis sur une démarche fortement interdisciplinaire, à l'intersection entre psychologie cognitive et philosophie, avec une préoccupation d'applications pour le développement de dispositifs de substitution sensorielle. Le premier axe de mes recherches concerne les dispositifs de substitution sensorielle qui visent à assister ou à remplacer une ou plusieurs fonctions d'un organe sensoriel défaillant à l'aide d'un autre organe sensoriel. Mes travaux scientifiques s'inscrivent à la fois en amont et en aval de la conception des dispositifs. Ils impliquent l'étude de la nature et des possibilités offertes par la substitution sensorielle, des facteurs centraux à leur apprentissage, de leur adéquation à la population cible en prenant en compte les spécificités individuelles et de la possibilité d'améliorer leur conception par l'utilisation des correspondances intermodales. J'ai aussi collaboré à la mise au point et aux tests de trois de ces dispositifs. Le second axe des mes recherches se penche sur la construction multisensorielle de nos représentations spatiales en fonction des modalités sensorielles impliquées. Par exemple la vision se base sur un référentiel centré sur les yeux, l'audition sur un référentiel centré sur la tête et le toucher à la fois un cadre de référence centré sur la surface du corps et un centré sur la tête. Mon objectif est de comprendre comment ces cadres de référence interagissent pour donner lieu à une représentation fonctionnelle de l'environnement et comment ils se modifient lors de l'utilisation de dispositifs de substitution sensorielle, et plus généralement d'outils. Pris ensemble, ces deux axes de recherche visent à un niveau fondamental à éclairer la question de la plasticité des représentations de nos espaces corporels et distants. A un niveau appliqué, ils visent à l'amélioration de la conception et de l'apprentissage des dispositifs de substitution sensorielle.

## **I. General introduction**

In this manuscript, I present the scientific research conducted since my PhD, in 2004. The presented studies were carried out during a year I've spent as ATER (associate lecturer) at UTC, Compiègne University (2004-2005), during three years as a post-doctoral researcher (2005-2008) at Oxford University (under the supervision of Prof. Charles Spence), at the University of Antwerp (under the supervision of Prof. Erik Myin), and at the Laboratory Neurophysics and Physiology, Paris (CNRS UMR 8119, under the supervision of Agnès Roby-Brami and Sylvain Hannequin). The studies were then conducted after I was hired by the CNRS, first at the LIMSI (CNRS UPR 3251), Orsay (2008-2014), then at the ISIR (CNRS UMR 7222), Paris, where I am since January 2015.

My research focuses on perceptual processes and in particular on the relations between sensory modalities, both in their correspondences and in their interactions. They were conducted thanks to national collaborations (ISIR, UMR 7222; IJN, UMR 8129; ISM, UMR 7287; LPP, UMR 8158; LNP, UMR 8119; IRCAM, UMR 9912; Impact, U864; Costech-UTC; LIMSI, UPR 3251; Institut de Recherche Paul Bocuse), international collaborations (Department of Experimental Psychology, Oxford University; Institute of Philosophy, London University; Università degli Studi di Milano-Bicocca, Milan; Department of Cognitive Neuroscience, University of Bielefeld; Antwerp University; York University), and collaborations with the industry (Caylar, Ile de France; Pernod-Ricard, Paris; Roger McMahon & Associates, Glasgow; Oculus Research).

The conducted research is based on an intrinsic multidisciplinary approach. The tested hypotheses come from a joint consideration of theoretical models, experimental and clinical approaches, with a preoccupation of applied outputs in the field of perceptual rehabilitation. My scientific work gave rise to 34 journal articles, 12 book chapters, 68 communications in international conferences with proceedings, and 60 communications without proceedings (1921 quotes, according to Google Scholar, for an h-index of 19). My research on sensory substitution allowed me to participate to TV and radio documentaries, as well as general public articles.



## **II. Sensory substitution**

The first and main axis of my research concerns sensory substitution devices. These devices aim at compensating one or several functions of a deficient sensory modality by converting stimuli that are normally accessed through this deficient sensory modality into stimuli accessible by another sensory modality. For instance, these devices can convert visual information into sounds or tactile stimuli. Sensory substitution devices have broad applications such as sensory rehabilitation and perceptual augmentation. For instance, visual-to-tactile (e.g., Bach-y-Rita et al., 1969; White et al., 1970) and visual-to-auditory (e.g., Hanneton et al., 2010; Meijer, 1992) conversion systems were designed in order to assist blind and visually impaired people. In most visual-to-tactile devices, the translation code is analogical. For instance, a visual circle can be translated into a circular pattern of tactile stimuli. Non-analogical codes have also been used, for instance by converting distance into vibrations (e.g., Farcy et al., 2006; Maidenbaum et al., 2014). The code used in visual-to-auditory devices translates several dimensions of the visual signal into dimensions of the auditory signal. For instance, the vOICe (Meijer, 1992) translates vertical position into frequency, horizontal position into time scanning, and visual brightness into auditory loudness. Other visual-to-auditory devices convert colour information into musical notes (e.g., the EyeMusic, Levy-Tzedek et al., 2014) or use different orchestral instruments (the See CoLoR; Bologna et al., 2009). Numerous devices were designed to compensate other deficits, such as proprioceptive deficits (e.g. see Danilov et al., 2007; Diot et al., 2014). In the field of perceptual augmentation, these devices were developed in order to gain access to visual information, either under degraded conditions of perception (such as fire fighting or military operations, e.g. see Jones et al., 2009) or when the amount of visual information is thought to be too important (such as in car driving; e.g. Ho & Spence, 2008). My research focused mainly on visual-to-tactile and visual-to-auditory sensory substitution devices that were designed to compensate for visual impairments.

Studies conducted on sensory substitution devices revealed an important structural and functional plasticity of the central nervous system. Regarding structural plasticity, studies have shown that the use and practice of visual-to-tactile (Ptito et al., 2005; see also Kupers et al., 2006) and visual-to-auditory (e.g., Striem-Amit & Amedi, 2014) devices result in increased activation in blind people's visual cortex (for reviews, see Proulx et al., 2014; Stiles & Shimojo, 2015; see also Ptito et al., in press, for a review and discussion of the functional correlates of increased activation in the visual cortex). Regarding functional plasticity, a variety of devices have been tested across a variety of tasks. In particular, studies have revealed that these devices allow their users to perform localisation tasks (e.g., Janson, 1983; Levy-Tzedek et al., 2012; Proulx et al., 2008) and simple form, as well as complex shapes recognition (e.g., Arno et al., 2001; Auvray, Hanneton, & O'Regan, 2007). In addition, users of sensory substitution devices can build, to some extent, a perceptual

space that possesses visual characteristics although the sensory inputs come from another sensory modality (e.g., Renier et al., 2005).

Since their inception in the sixties, sensory substitution devices proved to be fruitful tools to investigate questions in the fields of philosophy of perception, experimental psychology, and studies of brain plasticity. However, in spite of their potential and the perspectives they open, sensory substitution devices remain under-used to date. Thus, one of the major challenges for scientists working in the field of rehabilitation is to favour the use of aid devices by impaired people. The limited success of sensory substitution devices can be explained by insufficiencies in the analyses of what is sensory substitution, insufficiencies in the understanding of impaired people's perceptual specificities, and deficiencies in the design of the devices. To overcome these insufficiencies my research lies both upstream and downstream the design of sensory substitution devices. In particular, my research involves a conceptual understanding of the nature and possibilities offered by sensory substitution devices, studies of the central factors underlying their learning, their adequacy to the target population, taking into account their individual specificities, and the possibility of improving their design by the use of crossmodal correspondences. I have also collaborated to the design and tests of three sensory substitution devices, two visual-to-auditory conversion systems and a visual-to-tactile one.

## **II.1. Conceptual understanding of sensory substitution**

I have investigated conceptually the question of the nature of the processes involved in the use of sensory substitution devices in collaboration with Ophelia Deroy (University of London), Erik Myin (University of Antwerp), Maurice Ptito (McGill University, Montreal), Ron Kupperts (University of Copenhagen), Laurence Harris (York University), Mirko Farina (University of Edinburgh), and Gabriel Arnold (post-doctoral researcher).

From a theoretical point of view, one of the main goals is to understand what underlies the acquisition of new recognition and localization skills that are usually characteristic of a given sensory modality by means of another. In particular, one important question raised is to which sensory modality does the perception with a sensory substitution device belong? Two opposite theses were, at first, put forward: The dominance thesis (e.g., Block, 2003; Humphrey, 1992; Prinz, 2006) according to which perception with a sensory substitution device remains in the substituting modality (touch or audition) and the deference thesis (e.g., Hurley & Noë, 2003; Noë, 2004; O'Regan, 2011) according to which perception switches to the substituted modality (vision). The deference thesis has opened the door to over-optimistic claims involving the idea that users of visual-to-tactile substitution devices would become able to “see with the skin” (White et al., 1970) or to “see with the brain” (Bach-y-Rita et al., 2003). These optimistic claims have been echoed to the wider audience

including potential users of these devices: sensory substitution devices have been advertised as “rewiring brains to see with sound” and “restoring a form of sight to the blind” (Trivedi, 2010).

With my collaborators (Arnold, Pesnot-Lerousseau, & Auvray, 2017; Auvray & Farina, 2017; Auvray & Myin, 2008; Deroy & Auvray, 2012, 2014), we suggested that it is time to go beyond the dominance versus deference debate, and that perception with a sensory substitution device goes beyond assimilation to either the substituting or the substituted modality. According to this view, sensory substitution should be understood as being vertically integrated on pre-existing capacities that involve the substituting and the substituted sensory modalities, as well as cognitive processes (see figure 1).

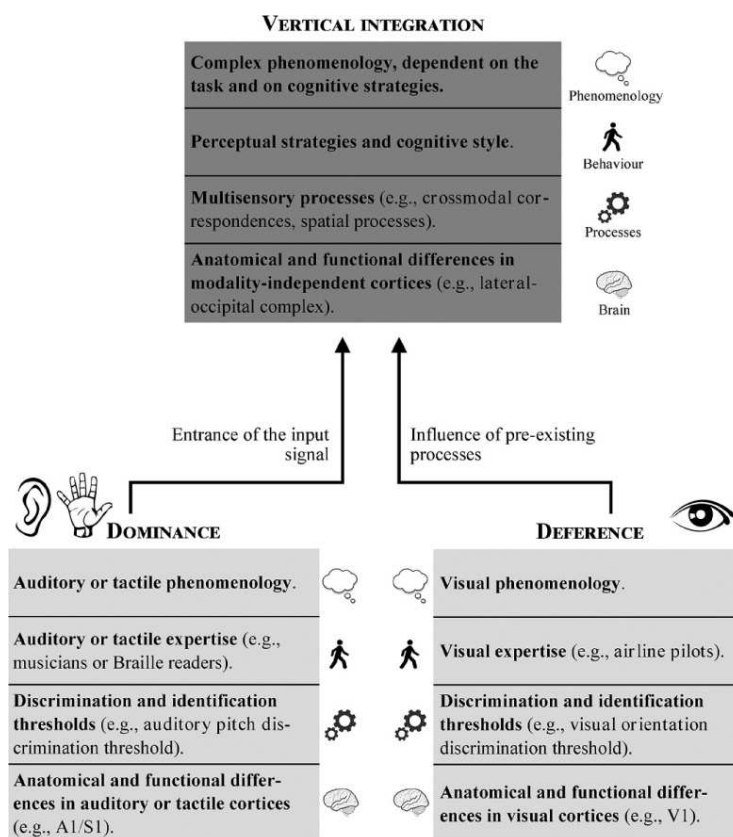


Figure 1. Three alternative views on the processes involved in sensory substitution, with their specificities at the behavioural, neural, and phenomenological levels. Illustrative examples are given into brackets (figure from Arnold, Pesnot-Lerousseau, & Auvray, 2017).

To go deeper into the defended view, the dominance versus deference debate remains based on an implicit perceptual equivalence which we suggest to identify as a perceptual assumption. The influence of this perceptual assumption is visible in the fact that researchers accept or target equivalences between using a sensory substitution device and the exercise of a sensory modality. In other words, the perceptual assumption considers that sensory substitution follows what occurs with canonical cases of perception through one of the

typical sensory modalities that is as specialized channels for transducing external information. To us, this perceptual assumption has led to a confirmation bias in the interpretation of the results. The results of existing studies have been filtered out of the negative evidence, or data fitting less well with this assumption, while the remaining evidence has been seen as confirming the equivalence between using a sensory substitution device and perceiving through one of the canonical senses. Furthermore, the experimental protocols themselves are built with the perceptual assumption in mind which, in turn, constrains or limits the kind of data that can be gathered. In the series of articles summarized below, we sought to stress the limits of the perceptual models and the perspectives that can be opened by alternative models.

First, in collaboration with Erik Myin (Auvray & Myin, 2009), we discussed the dominance versus deference debate. We had recourse to the criteria that have been used for the taxonomy of our sensory modalities to determine which of them are fulfilled by the use of sensory substitution devices. These criteria involve sensory organ, stimuli, properties, qualitative experience, behavioural equivalence, dedication, and sensorimotor equivalence. The reviewed criteria did not allow favouring the dominance or the deference theses, and most of them appeared as a matter of degree or a matter of interpretation. We subsequently argued that the evidence leads to an alternative view, according to which the experience after sensory substitution is a transformation, extension, or augmentation of our capacities rather than being equivalent or reducible to an already existing sensory modality. We developed this view by comparing sensory substitution devices to other “mind enhancing tools” such as pen and paper, sketchpads or calculators. In particular we built on the transformative view of mind enhancing tools (e.g., Clark, 2003; Menary, 2006, 2007) according to which they do transform cognition in a qualitative way. In this sense, these tools not only facilitate established cognitive processes, they can also allow for the appearance of novel cognitive operations, which simply would have been impossible without them. For example, without the proper means and tools to write down, calculate, or draw diagrams, human cognitive abilities would not have evolved to their current state (Van Leeuwen, Verstijnen, & Hekkert, 1999). To us, an analysis of sensory substitution in terms of mind enhancing tools unveils it as thoroughly transforming sensory experience and as giving rise to a novel form of interaction with the environment.

In a second step, we refined this view with Ophelia Deroy (see Deroy & Auvray, 2012, 2014), proposing a vertical integration model. We argued that learning to use a sensory substitution device should not be thought of as occurring horizontally. Indeed, the use of sensory substitution devices does not have to fit within the concept of a sense even if it apparently serves similar functions (e.g., identification and localization). The corresponding experience does not fit among the sensory modalities and it requires the existence of (some

of) these modalities. Experience with sensory substitution devices is built up from existing sensory modalities both in terms of the used receptors and of the invented code the devices rely on. In addition, this relation is of interdependence or crafting. In this sense, the new skill starts from an existing one, and although it becomes progressively more independent, it does not become totally detached from the initial elements. We used an analogy with dual-route models of reading (e.g., Coltheart & Rastle, 1994), according to which, learning to use a sensory substitution device is no longer to be thought of as being merely a matter of perceptual learning or adaptation, but as the building of a parallel access to cognitive and spatial representations that get grafted onto some pre-existing perceptual-cognitive route (e.g., sounds to objects and spatial representations in the case of visual-to-auditory conversion systems). What sensory substitution reveals in this case, is not strictly sensory plasticity nor perceptual emergence, but mostly culturally driven multisensory plasticity, that is the margin left for exploiting and redirecting the existing rules of multisensory and crossmodal interactions to build new cognitive routes between existing components. To summarize, here sensory substitution is treated as involving both perceptual and cognitive processes, which take into account integration of the novel information with the existing perceptual-semantic route.

In addition, with Mirko Farina (see Auvray & Farina, 2017), we discussed the claims made that long-term use of sensory substitution devices may induce, in practiced users, the emergence of forms of synaesthesia (e.g., Proulx, 2010; Ward & Meijer, 2010; Ward & Wright, 2014). To clarify this position, we proposed to dissociate the substituted information from the associated phenomenology and to verify for each, if they fulfil the criteria used to define congenital synaesthesia (see Table 1). Our analysis highlighted that, when the concurrent is considered to be the associated phenomenology there is a crucial lack of empirical data, as the hypothesis arises from verbal reports of only two users. These preliminary reports suggest that there might be a consciously perceived concurrent. There can be idiosyncrasy, although narrower than in congenital synesthesia, and there are no substantial data to know if there can be automaticity and consistency; although these cannot be ruled out either. When the concurrent is considered to be the substituted information, the existence of an inducer concurrent pairing can be acknowledged. However, the extent to which both the concurrent and the inducer are consciously perceived awaits further empirical data. In addition the concurrent is linked to the inducer more in a relative way (for instance the louder, the brighter) than in an absolute way (which is closer to cross-modal correspondences than to synesthesia; see Deroy & Spence, 2013) and there is no idiosyncrasy. In summary, sensory substitution can hardly be claimed to match the criteria established to single out genuine forms of synesthesia. As a consequence, anyone defending SSD-use as being synesthetic will have to give an account of why one or several criteria to

define synesthesia can be loosened in some cases without weakening the understanding and definitions given to congenital synesthesia.

Cases	Criteria	Inducer-concurrent pairing	Idiosyncrasy	Automaticity of the process	Consistency over time
Congenital synesthesia		Yes	Yes	Yes	Yes
SSD: associated phenomenology		Debated	Yes (narrow set)	Lack of Data	Lack of Data
SSD: substituted information		Lack of Data	No	Lack of Data	Yes
Crossmodal correspondences		No	No	Yes	Yes

Table 1. Summary of the criteria used to define synesthesia, which are fulfilled by the use of sensory substitution devices. In the table we also added congenital synesthesia and crossmodal correspondences. The terms yes and no are used when the claim is not controversial, debated is used when there are existing data, but their interpretation is subject to controversy, lack of data is used when more empirical data are needed.

Finally, with Gabriel Arnold (post-doctoral researcher) and Jacques Pesnot-Lerousseau (master student), we built on the vertical integration model in order to investigate the predictions that can be made from this model (Arnold, Pesnot-Lerousseau, & Auvray, 2017). In particular, the vertical integration model involves the idea that, at the perceptual level, performance and experience with the device will depend on each person's abilities. Second, if the use of sensory substitution devices is both perceptual and cognitive, then this leaves room for perceptual strategies that will differ as a function of people's specificities and cognitive style. In other words, performance and experience with the device will then be done with different weights attributed to each of the sensory modalities. These weights will depend on people's individual capacities, perceptual preferences, and cognitive styles. Note that these weights might subsequently vary as a function of expertise, mastery of the device, and type of task. To give an example, it might be the case that a musician will be able to perform auditory analyses that non-musicians would not, consequently giving more weights to auditory processes, and possibly having a predominant auditory experience when using the device.

## II.2. Seeing the thunder while still hearing it

In order to start and investigate behaviourally the vertical integration hypothesis, we conducted an experiment in collaboration with Gabriel Arnold (post-doctoral researcher) and Jacques Pesnot-Lerousseau (master student). **Note that this study is included as the first of the selected articles in the Annex (Pesnot-Lerousseau, Arnold, & Auvray, submitted).**

In this study, blindfolded sighted participants were trained with the visual-to-auditory conversion device The vOICe (Meijer, 1992), which translates visual images into soundscapes (see Figure 2A). After training, with the blindfold removed, participants were tested to see whether auditory stimuli would spontaneously evoke visual images. To evaluate such potential involvement of visual processes, we took advantage of a Stroop-like paradigm (Stroop, 1935). In the original Stroop task, people are requested to name the colour of coloured words. When the meaning of the word is also a colour (e.g. the word “red” written in blue), it interferes with the participant’s recognition of the word’s colour. For instance, it takes longer to say that the colour of the presented word is blue when the written word reads “red” rather than “blue”. Thus, even if the task consists in naming the colour of a word, reading processes and the ensuing access to the words’ meaning are automatically triggered.

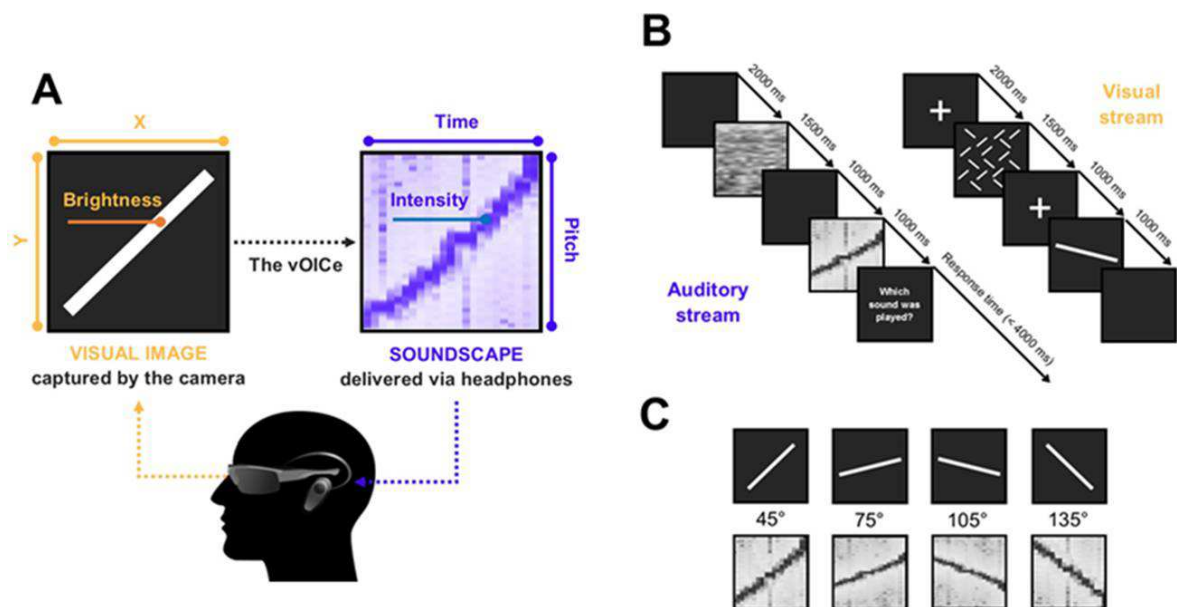


Figure 2. Illustration of the methods. (A) The visual-to-auditory conversion device (The vOICe). The images recorded by the video camera embedded in the glasses are converted in real-time into soundscapes that are presented to the participants via headphones. The device converts the vertical dimension into auditory frequency, the horizontal dimension into time scanning, and visual brightness into auditory loudness. (B) Description of the crossmodal interference task. (C) The auditory targets (spectrogram) and visual distractors that are used in the crossmodal interference task.

The same rationale was used in our study. The participants completed a crossmodal interference task both before and after a 3-hour training session with The vOICe, which involved object identification, object localization, and orientation discrimination tasks. The crossmodal interference task consisted of a sound recognition task with a simultaneous presentation of visual images (see figure 2B). The sounds corresponded to the auditory conversion of visual lines with The vOICe. Visual distractors, presented simultaneously,

consisted in these abovementioned visual lines; some had the same orientation as the sound, whereas others presented a different one (see figure 2C). After training with the visual-to-auditory conversion device, if visual processes are automatically triggered upon hearing the sounds, i.e. if auditory stimuli are mentally converted into visual images, then the presentation of visual distractors should interfere with the participants' performance in the sound recognition task. Should this be the case, their performance ought to be facilitated when the auditory target and visual distractors have the same orientation, and disturbed when auditory target and visual distractors have different orientations. Conversely, if auditory stimuli are processed solely as sounds, i.e. without any involvement of visual processes, then visual distractors should not affect the participants' performance. In addition, in order to investigate the possible influence of auditory processes, the participants completed low-level auditory tests before and after training. Furthermore, prior training, and after training in each of the tasks, they were questioned about their subjective experience by means of Lickert scales (see figure 3, for the different steps of the procedure).

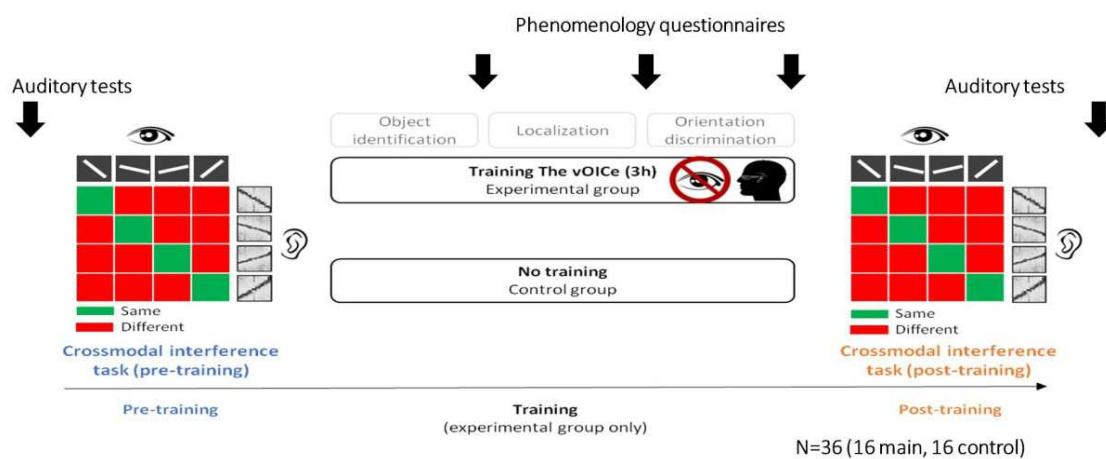


Figure 3. The different steps of the experimental procedure.

The first and main result from our study, obtained in the crossmodal interference task, showed that, before training, the visual images did not influence the participants' responses. After training, however, they interfered with the auditory recognition task. In particular, they disturbed the participants' responses when the auditory soundscape did not correspond to the conversion of the visual image. This visual interference effect reveals a rapid functional plasticity, as users, once trained, can associate visual imagery to auditory stimuli (see figure 4).

Second, the participants' performance during training for localisation and recognition tasks, depended on their auditory abilities. The correlations between the participants' auditory scores and their performance with the device revealed that those participants with higher auditory abilities performed better during the training tasks with the device than those with lower abilities. Third, the participants' associated phenomenology differed as a function of



the task, as in Auvray *et al.*'s (2007) study. The overall phenomenology remained auditory, which is likely due to the fact that the stimuli consisted in sounds. However, in the object recognition task, the participants had additionally an important visual experience; whereas in the object localization task they had an additional sonar-like experience, that is, an experience mainly based on spatial information of localization or distance from the body, without necessarily forming visual images of the object to localize. Another interesting result is that such difference in the phenomenology as a function of the task – which probably results from the fact that several different processes are involved – is stronger in participants with good auditory capacities than in participants with low ones. It is likely that good auditory capacities allow the participants to identify those dimensions of the auditory signal that are the most relevant to the task while discarding the less relevant ones. Participants with low auditory capacities have a phenomenology which did not differ with the task, which suggests a lower plasticity in the involved processes. As their auditory inputs are poorer, it could be the case that they are less able to disentangle the different auditory dimensions and consequently to analyse them independently.

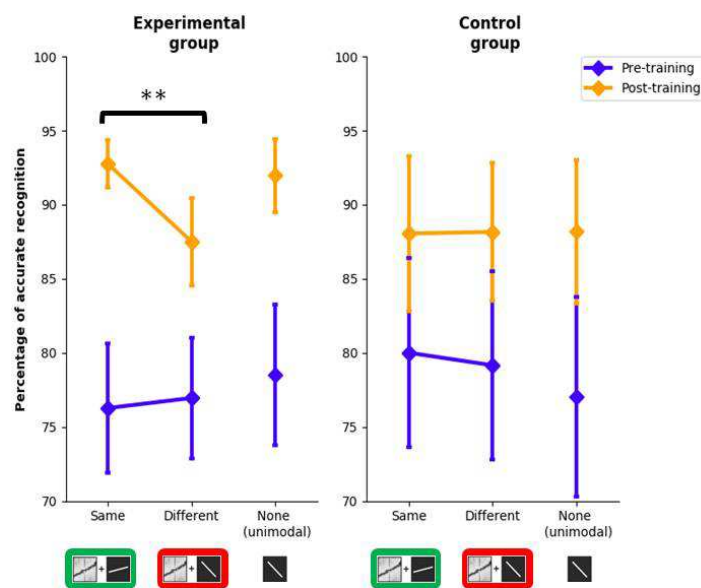


figure 4. Mean accuracy (percentage of correct recognition) in the crossmodal interference task across participants of the experimental and control groups, as a function of the Session (pre-training, post-training) and the Type of visual distractor (same, different, none). In the control group, pre-training and post-training sessions corresponded to the first and second sessions respectively, however without training with the device in between. Error bars represent standard error of the mean.  $**p < .01$ .

To summarize, the results of this study reveal that using a visual-to-auditory sensory substitution device induces plasticity at the functional and phenomenological levels. The association of the visual interference effect with the role of individual auditory abilities underlines the fact that functional plasticity is complex, and based on a multisensory architecture involving both visual and auditory processes. Our results show that, after

training, people visualize auditory stimuli, while still processing it auditorily. These results are in line with the vertical integration thesis, which posits that using a conversion device involves flexible behavioural processes and phenomenologies, with different weights attributed to the sensory modalities as a function of task, context, expertise, and individual differences (Arnold, Pesnot-Lerousseau, & Auvray, 2017). Learning to use a conversion device thus relies on pre-existing perceptual capacities, and it can involve different perceptual strategies, as a function of these pre-existing capacities. Note that these results are also in line with the metamodal inverse hierarchy model (Proulx et al., 2014, 2016) which proposes that the functional plasticity at stake occurs at higher levels, involving supramodal representations of shapes with a spatial integration of visual and auditory stimuli. More generally, our results reveal that sensory plasticity in humans is a complex phenomenon which depends both on the kind of processes that are involved and on individual specificities. To put it in William James' (1890) words, when using the visual-to-auditory sensory substitution system, people become able to see the thunder while keeping on hearing it.

Several questions remain open to scientific investigation. First, the question of whether the interference appears at the visual or at a supramodal level remains to be investigated. Second, we still need data on the questions of the influence of visual capacities on performance with the visual-to-auditory conversion system and the role of visual imagery in the processes (i.e., if we can obtain a crossmodal interference effect the other way round). We are currently running the experiments corresponding to this second question.

### **II.3. Sensory substitution and learning**

An important feature of the mastery of sensory substitution devices is their users' facility to learn them and the amount of training they require to do so. For instance, with the TVSS, an extensive training is required. Most users are able, in one or two hours, to explore their environment, approximate objects' positions, and describe their raw shape. However, the learning required to reach a more complete level of performance is estimated at around 8 hours with visual-to-tactile devices (Kaczmarek & Haase, 2003) and 10 to 15 with visual-to-auditory devices (Auvray et al., 2007), and an even a longer time to perform fine grained tasks, such as recognizing body postures (Striem-Amit & Amedi, 2014). It is thus crucial to investigate users' learning abilities, to identify the most efficient learning methods and those that are adapted to different users' profile. A crucial characteristic of learning lies in the ability to generalize, that is, the ability to extend the acquired perceptual abilities to both new stimuli and new perceptual conditions. With Gabriel Arnold (post-doctoral researcher), we focused for this part on visual-to-tactile sensory substitution devices. We have conducted behavioural research to investigate the extent to which tactile learning transfers across body

surfaces and the extent to which learning transfers to novel stimuli and to novel perceptual conditions.

### Perceptual learning 1. Tactile letter recognition transfers across body surfaces

In order to investigate the extent to which tactile learning transfers across body surfaces, we conducted an experiment in collaboration with Gabriel Arnold (post-doctoral researcher). **Note that this study is included as the second of the selected articles in the Annex (Arnold, & Auvray, 2014).**

One important claim of Bach-y-Rita is that, once trained, users of visual-to-tactile sensory substitution devices no longer feel the stimulation on their skin, where it occurs, but they directly attribute the stimulation as resulting from an external object, i.e. located at a distance (Bach-y-Rita & Kerckel, 2003). Bach-y-Rita then asserted that, consequent to this externalization process, the tactile stimulator array can be moved from one body surface to another, without loss in spatial localization abilities or other perceptual capacity. This claim, although recurrent, was only based on users' verbal reports and it has never been investigated with objective methods. Moreover, studies in the field of tactile perceptual learning suggest that performance improvement transfers only to body surfaces that are closely represented in the somatosensory cortex. These studies have however mainly used discrimination tasks of stimuli varying along only one feature (e.g., orientation of gratings) whereas, in sensory substitution, tactile information consists in more complex stimuli.

Our study aimed at investigating the extent to which there is a transfer of tactile learning in the recognition of high-level symbols (i.e., made of the combinations of several features rather than consisting of single features). To do so, a tactile letter recognition task was used. Letters were drawn on the participants' body by means of sequential vibrotactile stimulations. The participants first completed a baseline session in which they had to recognize tactile letters drawn on three different body surfaces: the belly, the front of the right thigh, and the right shin (see figure 5). They then underwent a training session in which the letters had to be recognized in only one of the three abovementioned body surfaces. Finally, the participants performed a post-training session on all three possible body surfaces. The amount of tactile learning was evaluated by computing performance improvement (both accuracy and response times) between the baseline and the post-training sessions. If there is a transfer of tactile learning from a trained to an untrained body surface, then the amount of performance improvement should be similar for trained and untrained surfaces. On the other hand, if tactile learning is specific to a given body surface, then performance improvement should be greater for trained than for untrained body surfaces. In addition, a control group of participants performed the baseline and post-training sessions, but without the training session in between. For these participants, performance is expected

to be either similar between the two sessions, or to improve in the second one but less than for the participants in the trained group.

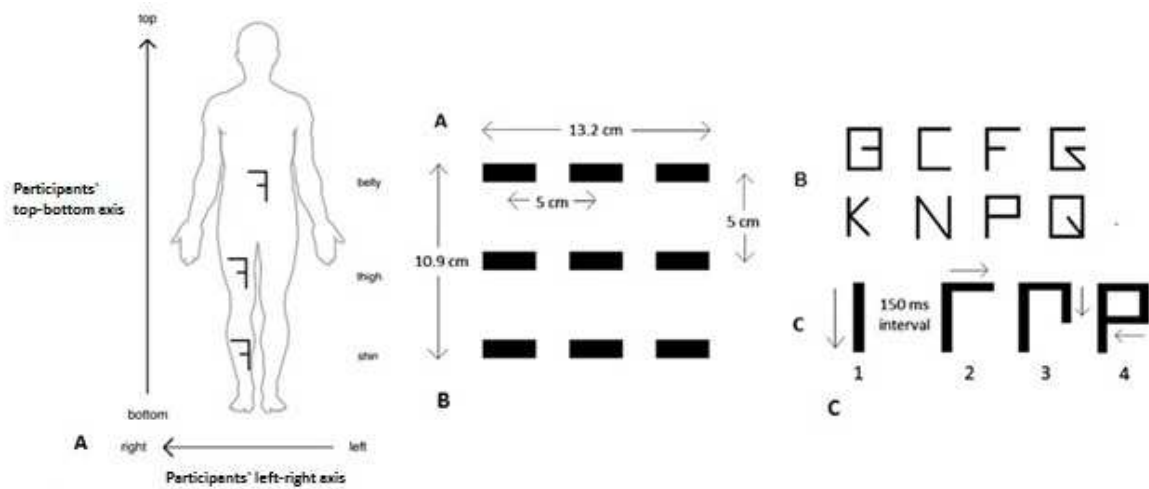


Figure 5 (A) Example of the letter F drawn on the three body surfaces (belly, thigh, and shin). The letters' left-right and top-bottom axes were always congruent to the participants' left-right and top-bottom axes. (B) Schematic figure illustrating the 3 × 3 array of rectangular vibrators. (C) The 8 letters used as tactile stimuli.

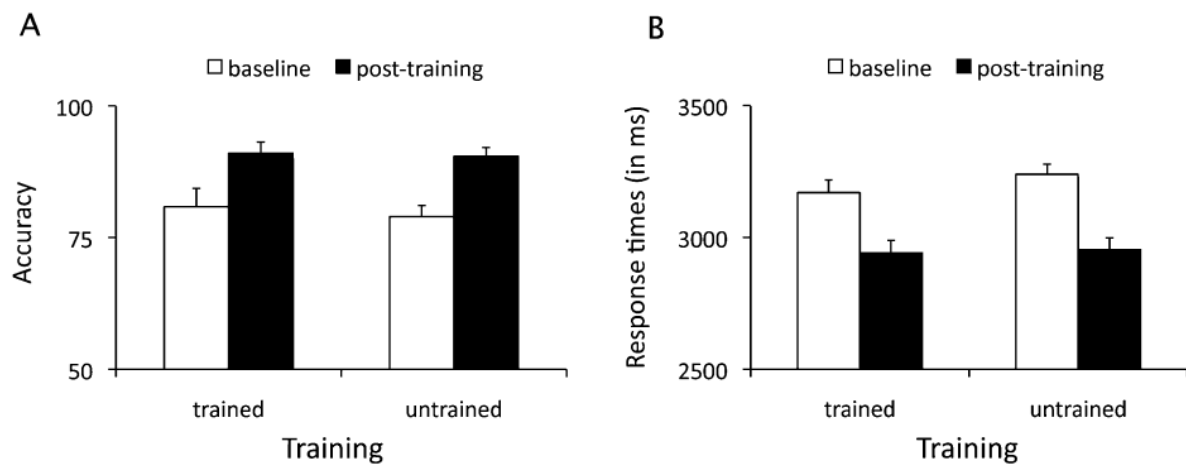


Figure 6. Participants' (A) accuracy and (B) response times for the trained and untrained body surfaces obtained during the baseline session (in white) and the post-training session (in black). Error bars represent the standard errors of the means.

The results of our study revealed that training improved both participants' accuracy and response latency (see figure 6). In addition, participants performing intensive training with trial-by-trial feedback (trained group) improved performance to a greater extent than the participants merely repeating the task twice on each body surface (control group). Second, performance improvement was not restricted to the trained body surface but it transferred to the untrained ones. Third, the obtained transfer of tactile learning was similar for surfaces represented in adjacent (e.g., belly and thigh) and in non-adjacent (e.g., belly and shin) areas

of the somatosensory cortex (see Merzenich et al., 1978; Nakamura et al., 1998, for descriptions of the somatosensory cortex's topographic organization). Taken together, these results reveal that there is a transfer of learning in tactile letter recognition, which occurs independently of the distance between body surfaces. These three results provide support for Bach-y-Rita's claim that training in sensory substitution improves perceptual abilities, and that these perceptual abilities transfer from one body surface to another, resulting in a relative independence from the stimulated body surface (Bach-y-Rita & Kercel, 2003).

These results have implications for the use of sensory substitution devices and vibrotactile systems by visually impaired people. First, they suggest that training on one body surface is beneficial to the entire body. As a consequence, users do not need to undergo an extensive training on all body surfaces. This study also provides better understanding of the learning process, which is a crucial feature for efficient use of tactile devices. It reveals that transfer of learning for tactile letters occurs within a relatively short time frame, i.e., the 90 min of our experiment. The longer learning reported for recognition of more complex objects with sensory substitution devices (Kaczmarek & Haase, 2003) indicates that the amount of training required to reach accurate object recognition with tactile devices depends on the complexity of the information that is provided.

#### Perceptual learning 2. Generalization to novel stimuli and novel perceptual conditions, specificity versus generalization of tactile learning

In order to investigate the extent to which tactile learning transfers to novel stimuli and to novel perceptual conditions, we conducted an experiment in collaboration with Gabriel Arnold (post-doctoral researcher). **Note that this study is included as the third of the selected articles in the Annex (Arnold, & Auvray, 2018).**

In this study, we more specifically investigated the perceptual learning of tactile alphanumerical stimuli. The learning protocol consisted in alternating a repeated list of symbols with lists of new symbols. The first experiment investigated the role of stimulus variability in the ability to generalize to new symbols during the learning of sets of four symbols. The results showed that recognition performance improved over time only for the repeated list. This result suggests that learning a small set of stimuli involves stimulus-specific learning strategies, preventing generalization. A second experiment revealed that increasing to six the set of learned stimuli results in higher generalization abilities. This result can be explained by greater difficulties in using stimulus-specific strategies in this case, thereby favouring the use of generalization strategies. Feature variability also appeared to be important to achieve generalization. Thus, as in visual perceptual learning, the involvement of stimulus-specific versus general strategies depends on task difficulty and

feature variability. A third experiment highlighted that tactile perceptual learning generalizes to changes in orientation.

These different results for small and larger sets of stimuli can be accounted for by the reverse hierarchy theory (RHT) of perceptual learning (Ahissar & Hochstein, 2004). According to this theory, the degree of specificity depends on the difficulty of the perceptual task and on the level of cortical processing that is required to perform the task. For a difficult task, involving low-level cortical processes that are highly specific to the characteristics of the task (e.g., fine discrimination of visual orientations), perceptual learning is specific to the characteristics of the task (e.g., the visual orientations that have been discriminated) and it does not generalize to new characteristics (e.g., new orientations). On the other hand, for an easier task, involving higher-level processing, perceptual learning is less specific to the low-level characteristics of the task or of the stimuli. Task difficulty, either during the learning phase or during the evaluation of generalization, has indeed been reported to influence both the specificity and generalization of learning (Jeter et al., 2009; Liu & Weinshall, 2000; Wang, Zhou, & Liu, 2013).

This study has applied implications for the design of learning protocols for sensory substitution. In particular, the results suggest that short training sessions with high variability and diversity will facilitate generalization and will allow for a more optimal use of the device in real-life conditions. However, one remaining question is how to achieve complete generalization, that is, how to rapidly gain optimal level of performance for entirely new objects. One of the possible cues is obviously the amount of training. Daily use of the device would allow users to become experts. Learning would also benefit from a highly diversified use of the device, with perception of different categories of complex objects and of different objects of the same category.

#### **II.4. Crossmodal correspondences for the optimization of the devices**

In this series of experiments, our aim was to use crossmodal correspondences in order to determine the most optimal codes for sensory substitution devices. This research was conducted in collaboration with Ophelia Deroy (University of London), Irène Fasiello, and Vincent Hayward (ISIR, UMR 7222). **Note that this study is included as the fourth of the selected articles in the Annex (Deroy, Fasiello, Hayward, & Auvray, 2016).**

One crucial question when designing sensory substitution devices is how to code at best visual dimensions into tactile or auditory dimensions. With visual-to-tactile generalist devices (i.e., excluding those coding only for distance, e.g., O'Brien, Auvray, Hayward, 2015), most of the time an analogical code is used, for instance, a visual circle gives rise to a circular pattern of vibrations (e.g., the TVSS, see White et al., 1970). With visual-to-

auditory devices, several dimensions of the visual information (most of the time horizontal and vertical dimensions, depth, brightness, and sometimes colour) are coded using several auditory dimensions (most of the time pitch, time scanning, inter-aural differences, loudness, and timber). The question arises as to how to determine the pairings between auditory and visual dimensions? The four main visual-to-auditory devices (e.g., the vOICE, Meijer, 1992; The PSVA, Capelle et al., 1998; Vibe, Hanneton, Auvray, & Durette, 2010; the EyeMusic, Abboud et al., 2014) have been designed by using the correspondence between direction of pitch and direction of movement (higher-higher; lower-lower). However they haven't tested whether this correspondence, intuitive for the designers and well-documented in audiovisual contexts for sighted individuals (see Parise & Spence, 2012; Spence, 2011) was also relevant for blind people.

In order to investigate the relevance of the pitch-elevation correspondence for blind persons, we had to test the audio-tactile analogue of this correspondence. Across four experiments, we tested the interactions between the direction of tactile movement (inward vs. outward movement on the fingertip) and changes in auditory frequency (increasing vs. decreasing pitch), see figure 7 for the description of the device used.

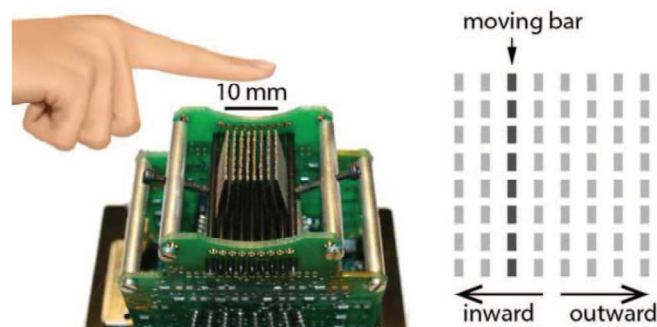


Figure 7. Tactile stimulator device and representation of the tactile stimulus.

A first experiment investigated this correspondence using a selective attention method, requiring participants to focus on tactile or auditory signals while ignoring the other one presented simultaneously. The results with this method did not reveal any significant compatibility effect, suggesting that the effect does not occur at an explicit level. In order to investigate if this effect occurs at the implicit level, we used a variant of the implicit association task (IAT, e.g., Parise & Spence, 2012). In this variant of the IAT, blindfolded sighted participants were required to focus on one stimulus at a time, while the association was only at the level of response buttons. Four stimuli were sorted using only two response buttons, each of which refers to two of the four stimuli. The compatibility effects between intuitively congruent pairs of stimuli (i.e., outward tactile movement, going from the inside of the finger toward the fingertip and increasing pitch, or inward tactile movement and

decreasing pitch) and incongruent pairs stimuli (i.e., the reverse associations) were measured. The results revealed a significant congruency effect between changes in pitch and changes in direction of tactile movement. Participants were for instance faster in responding to a tone increasing in pitch when the response button was shared with the outward tactile movement than when it was shared with the inward tactile movement.

A third experiment showed that this effect was similar in the conditions where the arm was placed vertically and horizontally, that is whether or not the inward-onward tactile movement corresponded to the free movement of an object subjected to gravity. Thus, a rising pitch is associated to a tactile movement going toward the fingertip, and a descending pitch to a tactile movement going inward toward the palm, independently of the hand being oriented upward or horizontally. This result is consistent with the correspondence occurring in a hand-centred frame of reference. Finally, the fourth experiment investigated the correspondence in blind persons. The results revealed that neither early nor late blind persons were sensitive to it. In addition, there was no effect related to the age of onset. Thus, vision seems to be necessary to mediate the correspondence between direction of pitch and direction of tactile movement (see figure 8).

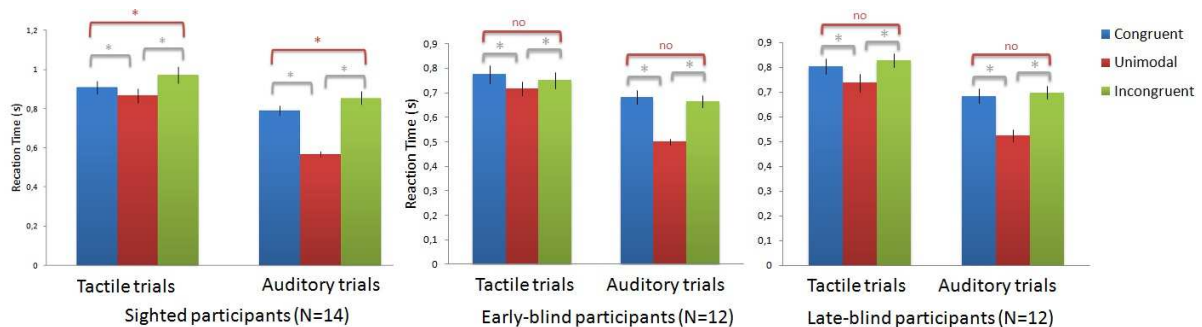


Figure 8. Participants' mean reaction times in Experiments 3 and 4, for the two conditions of stimulus presentation: tactile and auditory, and for the three congruency conditions: congruent, unimodal, and incongruent. Error bars represent the standard errors of the means. Note that, as there was no effect of arm position, in the graph, horizontal and vertical arm positions were collapsed.

The results of this study reveals that the audio-tactile correspondence between direction of pitch and direction of movement exists in sighted participants along both the horizontal and the vertical plane, but it is absent in early and late blind people. These results have methodological implications for the testing of crossmodal correspondences and for the design of sensory substitution devices. They indeed suggest that the IAT might be an interesting tool to pre-test the relevance of certain codings or combinations of auditory and tactile cues, to tailor them more specifically to their users. Our aim now is to investigate more systematically the possible crossmodal correspondences in audition, vision, and touch, focusing on the dimensions of the signals that are relevant for sensory substitution.



## **II.5. Prototypes of sensory substitution devices**

From an applied point of view, the challenges that are faced involve finding the best way to provide users with information usually allowed by vision as well as the best way to ergonomically adapt the devices to their users. I collaborated to the design and testing of three prototypes of sensory substitution devices, with the aim to target specific perceptual functions (localization, recognition, navigation) that can be impaired in blind people. I indeed believe that one of the reasons of the under-use of sensory substitution devices lies on the unrealistic promises of a generalist device that would restore the entirety of a deficient sensory modality. By contrast, blind persons more easily use technical devices devoted to a specific function (such as auditory or tactile software to navigate the internet).

The device Vibe. In collaboration with Sylvain Hanneton (LNP, UMR 8119), we have designed and tested, through a pointing task, a sensory substitution device that allows the real-time conversion of visual images into an audio stream (Hanneton, Auvray, & Durette, 2010). This device converts a video stream into a continuous stereophonic audio signal that conveys information coded from the video stream. The conversion from the video stream to the audio stream uses a kind of retina with receptive fields. Each receptive field drives a sound source and users listen to a sound that is a mixture of all these sound sources. Compared to other existing visual-to-auditory sensory substitution devices, Vibe is highly versatile in particular because it uses a set of configurable units working in parallel, for instance, the receptive fields of receptors can be large or not, identical or not, overlapping or not. In addition, The Vibe offers the possibility to enhance the binaural perception and differentiation by the listener of the sound sources in the plane of the picture by adding interaural disparity cues to the sounds like ITD (interaural time difference) or ILD (interaural level differences).

The NAVIG project (navigation aided by artificial vision and GNSS), funded by the French Agency ANR aimed at designing a navigation aid device for blind persons, using an auditory feedback. It involved GPS and artificial vision systems. The project was conducted in collaboration with three research laboratories (Irit, Cerco, and LIMSI), two companies (Spikenet Technology and Navocap), and an education center for blind persons (Institut des Jeunes Aveugles). At the LIMSI, in collaboration with Michel Denis (UPR 3251), Mathieu Gallay (post-doctoral researcher), and Lucie Brunet (master student) we focused more precisely on two research axes. First studying the processes underlying the representation of space that are specific to blind persons using navigation aid devices and a study specifying blind people's needs when navigating as well as whether those needs are satisfied by the prototypes designed in the project (Brunet, Darses, & Auvray, 2018; Gallay, Denis, & Auvray, 2013; see also Katz et al. 2012a&b, for the results obtained by the consortium).

Visuo-tactile conversion system. With the company Caylar and Gabriel Arnold (post-doctoral researcher), we received a funding from the CFI (Centre Francilien de l'Innovation), to develop a sensory substitution device involving 80 tactile vibrators, based on the tactile technology developed by the company. The aim was to design an interface highly configurable, so that each user can specify different options of the presentation of information, with the aim of gaining at best mastery of the device. For this device we had the chance to have three second year master students whose project was to evaluate the existing commercialized devices, the possibilities of intellectual property for these types of devices, and the industries involved in the project. After testing the first prototype, we agreed with the company on the adjustments to make and the final device is now developed and ready to be tested.

## **II.6. Conclusion on sensory substitution**

Although sensory substitution appears as a promising technology to compensate visual impairments, so far the developed devices remain barely used by visually impaired people, either because their design is not optimal, or because their use involves too many constraints and cognitive effort. Taking into account individual differences, as suggested by the vertical integration hypothesis, will allow an optimal appropriation of these devices. The learning protocols will also benefit from taking into account specific and generalized conditions of learning in order to favour the transfer of learning to novel stimuli and novel perceptual conditions.

In addition, identifying different perceptual strategies associated to different phenomenologies, with some people having more of a visual use of the device and other people having more of an auditory or tactile use, can have implications on the design of learning protocols. One possibility would be to respect each user's tendency and to customize the learning protocols as a function of the different types of use. For instance, for people having a visual use only, learning could be oriented toward integrating a substitution device so that the experience becomes increasingly visual. This can be done by favouring tasks designed to better apprehend depth, to learn how to form global images of the objects. For people having more tactile or auditory experiences, learning could be oriented toward favouring the processes specific to these modalities. For instance, learning to better discriminate auditory frequencies and auditory intensities will allow a better discrimination of the complex auditory stimuli coming from a visual-to-auditory device. Finally, the choice of the conversion codes will benefit from the identification of the crossmodal correspondences that may exist in the target population.

Regarding the conceptual aspects, the collected data helps making advances in the understanding of the processes involved when using a sensory substitution device. The

results suggest that performance with the device rely on users' individual sensory abilities and that experience takes roots both on the substituting and the substituted sensory modalities. The data collected on the subjective experiences, even if they remains too sparse to make strong claims, revealed that they are subject to inter- and intra- individual variability. The fact that there is dependence on both the task and individual abilities suggests that there is room for processes that are more cognitive, and thus more subject to individual variability. These results are in line with the view that sensory substitution devices are vertically integrated on pre-existing capacities that involve cognitive and perceptual processes belonging both to the substituting and to the substituted sensory modalities.

### **III- Multisensory spatial reference frames**

The second axis of my work investigates the multisensory groundings of our spatial representations, which vary as a function of the sensory modalities that are involved. This research axis was funded by a grant from the French agency ANR (FRESCO) and a post-doctoral funding grant (Fondation des Aveugles de Guerre). The research was conducted in collaboration with Gabriel Arnold, Jessica Hartcher-O'Brien (post-doctoral researchers), Charles Spence (Oxford University), Vincent Hayward (ISIR, UMR 7722), Jérôme Dokic, and Frédérique de Vignemont (IJN, UMR 8129).

#### **III.1. The variability in our reference frames**

For the story, at the beginning, this research axis was initiated to investigate the processes of distal attribution in sensory substitution, which will be further described below, in section III.8. The idea was to investigate how people, through training with a visual-to-tactile sensory substitution device, transitions from sensing the tactile stimuli on their body to perceiving the object as being located in front of them, in the 3-D space. One way to do so was to study the changes in the reference frames that they used before and after training. However, the results we started to obtain appeared far more complex than what we anticipated, and the more we dug into it the more it appeared fascinating, until becoming a research axis per se.

The impression of perceiving the external world from a single spatial origin, i.e., the self, requires a spatial unification of the multisensory information that comes to our body through the different sensory modalities (see Denis, 2018, for a review of the core aspects of spatial cognition). For instance, a fly is often perceived in a multisensory way: we can see it, we can hear the sound of its wings if its flying, and we can feel it when it lands on our body. However, we perceive not only three distinct stimuli but also one single unified object, the fly. Multisensory information is thus integrated in space and time, in order to perceive an external world that is spatially organized (Boyer et al., 2013; Meredith & Stein, 1986). Such

perception of a unified external world is a complex process because spatial reference frames are very different from one sensory modality to another. Visual information is initially coded in retinotopic coordinates<sup>1</sup>, whereas tactile information is initially coded in somatotopic coordinates, and auditory information is initially coded according to a head-centred reference frame. Spatial diversity is also present within one sensory modality. For instance, the spatial coordinates of tactile stimulation can be defined according to the stimulated body part, to the entire body, or to the external world. These different reference frames can conflict, as revealed by the longer time needed to report which hand, left or right, has been stimulated first when the arms are crossed as compared to uncrossed (Shore, Spry, & Spence, 2002). The additional time taken for localizing touch in the crossed-hand condition has been interpreted as reflecting a conflict between a body-centred and an environment-centred reference frame.

Regarding these spatial perspectives, the apparent primacy of self-centred perspectives has been challenged by studies revealing the extent to which people can adopt spontaneously self-centred versus other-centred perspectives (Arnold, Spence, & Auvray, 2016; Viaud-Delmon & Warusfel, 2014). On the one hand, as interoception is mainly body-centred, adopting a self-centred perspective on external stimuli is crucial for self-consciousness, as it allows for the integration of both interoceptive and exteroceptive information into a common egocentric body-centred reference frame (Blanke, 2012; Petkova, Khoshnevis, & Ehrsson, 2011). On the other hand, the adoption of a perspective that is decentred from the body can be advantageous in the case of perspective-taking, when it comes to interacting and communicating with others. As a consequence, it is important to understand what underlies the adoption of self-centred versus decentred perspectives and how they are integrated to give the impression of perceiving the external world from a single spatial origin.

In the experiments and theoretical work we conducted, we focused on investigating the spatial perspectives that are adopted in tactile perception. This is of particular interest to the study of bodily self-consciousness, especially when targeting the processes that integrate bodily stimulation with external stimuli (e.g., Aspell et al., 2013). Another specific characteristic of tactile perception is the spatial coincidence between the perceiver and the object of perception. Unlike visual stimuli, which are presented at some distance from the perceiver, tactile stimuli occur at the same location in space as the perceiver, i.e., the spatial location of the body. As a consequence, when interpreting tactile stimulation received on the

---

<sup>1</sup> More precisely, after this initial coding, visual information is then either maintained in retinotopic coordinates or it can be transformed into external, craniotopic, or motor coordinates depending on the task and goal of the observer, whereas for touch several coordinate systems already compete at early stages of somatosensory processing (Heed et al., 2015).

body surface, self-centred and decentred perspectives conflict with one another and there is a choice to be made between perceiving tactile stimulations from the body location (i.e. self-centred perspective) and from an external location (i.e. decentred perspective). It should be mentioned that these different self-centred perspectives can correspond to a mental localization of the tactile stimulation either on the body (proximal attribution) or on the external space (distal attribution, see figure 9).

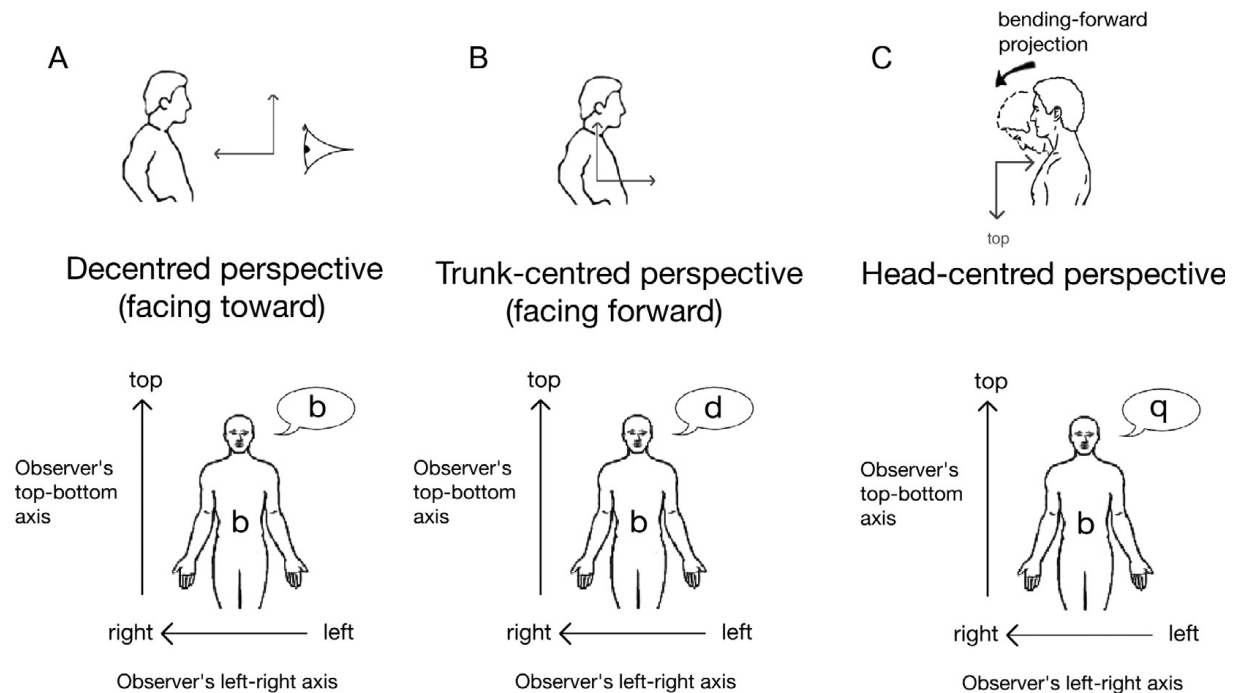


Figure 9: Illustration of the different spatial perspectives that can be adopted in the graphesthesia task, with the example of the tactile letter “b” drawn by the experimenter. (a) When the letter is drawn on the participant’s forehead, some participants perceive the letter “b”, assigning the top-bottom axis of the letter in the same direction as their own top-bottom axis, but the left-right axis in the direction opposite to their own left-right axis. These assignments may have resulted from a decentred perspective whose origin is located in front of the participant’s head. Other participants will perceive the mirror-reversed letter “d” instead, assigning the top-bottom and left-right axes in the same direction as their own body axes. These assignments may result from a self-centred perspective whose origin is located inside the head. (b) When the tactile letter is drawn below the head, here on the stomach, a third perspective centred on the head can be adopted, as if the head was bending forward to “see” the letter on the stomach. As a consequence of this bending-forward projection, the top-bottom axis of the symbol is assigned in the direction opposite to the participant’s own top-bottom axis and the participant perceives the 180°-rotated letter “q”.

In a review article (Arnold, Spence, & Auvray, 2017), we focused on the role of tactile spatial perspectives in self-consciousness. We review those studies that have used the graphesthesia task, that is, the task of recognizing ambiguous symbols (e.g., the letters b, d, p, and q) drawn on the body surface. The graphesthesia task provides an optimal paradigm with which to study the perspectives – self-centred and decentred – that are spontaneously adopted when interpreting bodily stimulations. Take the example of the lowercase letter “b” drawn on the forehead by an experimenter facing the participant (see figure 9). Recognizing

the letter “b” as drawn by the experimenter requires the participant to take a decentred perspective (i.e., decentred from the participant’s position and centred on that of the experimenter). However, because the lowercase letter “b” is ambiguous, it can also be recognized as the mirror-reversed letter “d”, from the participant’s self-centred perspective, as if the letter were to be projected forward and “seen” from the participant’s position. The crucial advantage of using these ambiguous symbols is that an intrinsic orientation of the symbol is impossible to determine and thus one specific perspective has to be taken in order to interpret the symbols. Consequently, when participants try to recognize these symbols, the pattern of their responses can be used to infer the way in which they have assigned the different axes to the stimulus and the perspective they have adopted.

Thus, the graphesthesia task can be used as a measure of self-localization in the body. For instance, in Ferrè et al.’s (2014) study, the participants had to recognize the letters b, d, p, and q, drawn on their forehead by the experimenter, while receiving galvanic stimulation of the vestibular system. The authors hypothesized that the vestibular system plays an important role in binding together multisensory information into an egocentric body-centred reference frame. When receiving galvanic stimulation of the vestibular system, participants adopted more of a self-centred perspective than a decentred one, compared to a sham stimulation in which a galvanic stimulation was provided on the neck’s skin, which provoked light tactile sensations without vestibular stimulation. This bias toward adopting a self-centred perspective with galvanic stimulation was interpreted as a reinforcement of the processes of the vestibular system, consisting in anchoring the self to the body (see also Deroualle, 2015; Pavlidou, Ferrè, & Lopez, 2018, on the influence of vestibular stimulation on the adoption of the egocentric reference frame). The graphesthesia task, as a measure of which perspectives are spontaneously adopted, can thus provide a reliable indicator of the potential disruption of the link between the self and the body, as people experiencing out-of-body experiences have been reported to adopt more spontaneously decentred perspectives than others (Blackmore, 1987).

The results from the literature also reveal that the adopted perspectives vary significantly with spatial factors (such as the location and orientation of the stimulated body parts), personal factors (such as gender, cognitive style, self-focused attention, spatial abilities), and on interpersonal factors (such as conflict or dominance in relationships, see Arnold, Spence, & Auvray, 2017, for the details of these variations). Such results suggest that the self can adopt a multiplicity of spatial locations. However, the unity of the self can partly be explained by the predominance of a head-centred perspective. Indeed, spatial information coming from the different senses can be integrated into one common and unified reference frame, usually an eye-centred reference frame, in the goal of performing actions (Cohen & Andersen, 2002). Spatial information coming from vision (Boussaoud & Bremmer, 1999),

audition (Zwiers, Van Opstal, & Paige, 2003), and touch (Harrar & Harris, 2010; Ho & Spence, 2007; Pritchett & Harris, 2011) has indeed been reported to be coded according to the direction of the eyes. The spatial transformation of multisensory information into a common eye-centred reference frame may contribute to the perception of a unified external world rather than multiple worlds. As a consequence, it can also induce the feeling of having one single unified self rather than multiple selves because the world is perceived from one single origin, i.e., from a self-centred perspective.

On the other hand, perspective-taking abilities contribute to the distinction between self and others, reinforcing self-consciousness. Indeed, if self-centred perspectives have primacy in our perception of the world, one can also adopt a perspective that is decentred from that of the self. A decentred perspective can correspond to the spatial viewpoint of another person (i.e., alter-centred, second-person, or third-person perspective), or simply centred at a different location that is not necessarily occupied by another person. Adopting different perspectives therefore facilitates imagining how the environment would be perceived from another point in space, and, importantly, understanding how the environment is perceived by others. The graphesthesia task thus appears as a promising tool to investigate the reference frames that can be adopted. From an experimental point of view, many questions remain to-be-investigated. We particularly focused on the questions of (III.2) whether the adopted perspective is natural or cognitive choice, (III.3) the influence of vision on the reference frames that are adopted, (III.4) the question of its embodied nature, (III.5) the role of proprioceptive information on spatial perspective taking, (III.6) the question of the transfer of a learned reference frame, (III.7) the specificities of the reference frames adopted on the hand, and (III.8), the processes of distal attribution.

### **III.2. Adopting a reference frame, a natural process or a mere cognitive choice?**

In order to investigate the extent to which adopting a given reference frame is a natural process or a mere cognitive choice, we conducted an experiment in collaboration with Gabriel Arnold (post-doctoral researcher) and Charles Spence (Oxford University). **Note that this study is included as the fifth of the selected articles in the Annex (Arnold, Spence, & Auvray, 2016).**

In this study, the participants were presented with the lowercases letters b, d, p, and q. These letters were presented by means of a matrix of 9 vibrators placed on the participants' stomach. At this location, three different perspectives can be adopted (see figure 10), a decentred perspective oriented toward the participant's stomach (response b); a trunk-centred perspective oriented forward the participant (response d); a head-centred perspective, as if the head was bending forward to "see" the tactile stimulation (response q).

The participants were requested to respond, for each stimulus, which of the four letters they perceived.

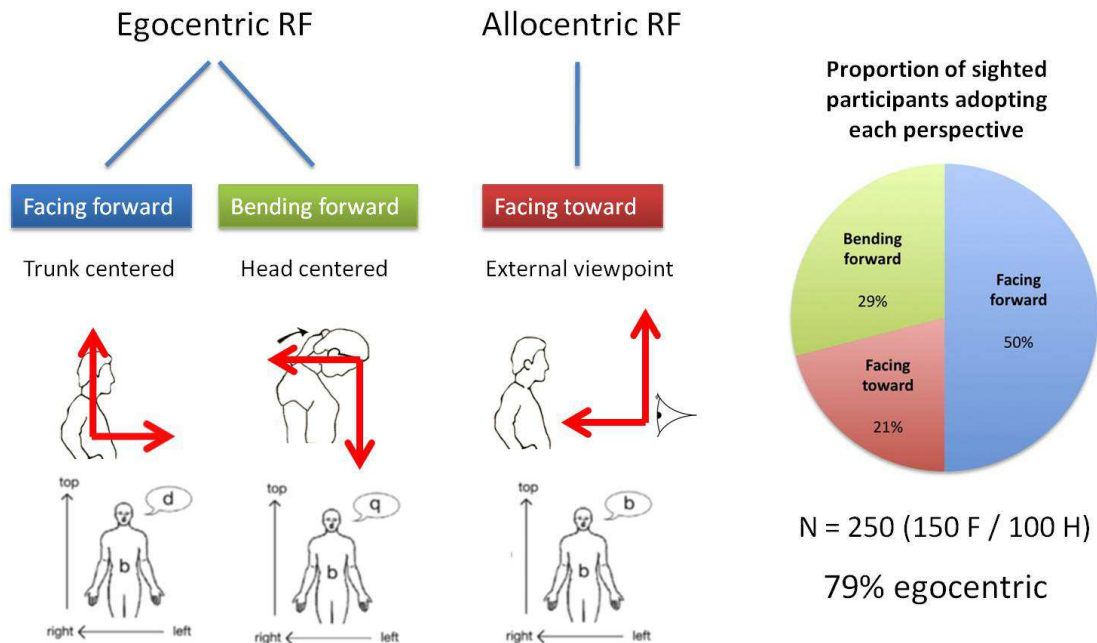


Figure 10. On the left, illustration of the three possible perspectives that participants can adopt when interpreting ambiguous symbols displayed on the body surface. On the right, proportion of the participants adopting each of the perspectives.

The experiment was composed of three sessions. In Session 1, the participants were free to adopt any perspective that they wanted to recognize letters. In Sessions 2 and 3, they were instructed to adopt a specific perspective. In one of these two sessions, the imposed perspective was the same as in Session 1. In the other session, the imposed perspective was different. Half of the participants performed Session 2 with the same perspective as in Session 1 and performed Session 3 with the different one. The other half did the opposite. These imposed perspectives allowed for the evaluation of any cost associated with adopting an unnatural perspective. If the perspective that is adopted freely is natural, then imposing a different perspective should produce a cost in terms of recognition performance. In addition, in order to evaluate whether the cost of adopting an unnatural perspective is simply explained by changes in perspectival instructions or by the difficulty that is associated with disengaging from a perspectival choice, we evaluated whether performance would improve when the participants returned to the natural perspective after adopting an unnatural one.

We first computed the proportions of the participants that adopted each of the perspectives. To do so, we included the 80 participants who performed the present experiment together with a further 170 participants who performed the same first session and were then included in other studies. On these 250 participants, 49.6% adopted the trunk-centred, 29.2% the



head-centred, and 21.2% the decentred perspectives. Thus, around 4 out of 5 participants adopted a self-centred perspective (centred either on their head or on their trunk), whereas only 1 out of 5 adopted a decentred perspective (see figure 10).

We then computed the participants' mean accuracy and response times in each of the three blocks (i.e., on the 80 participants of this study). Results from the ANOVA revealed that recognition performance decreases with an unnatural perspective and returns to its previous high level with a natural one (see figure 11). This result supports the view that the perspective adopted is natural. In addition, when a different perspective was imposed, the participants made more errors corresponding to the adoption of their natural perspective than other errors. Some of the observers perceive spatial relations better from a self-centred perspective whereas others perceive better from a decentred perspective. Interestingly, the results also revealed a greater cost for self-centred participants to adopt a decentred perspective than for decentred participants to adopt a self-centred perspective.

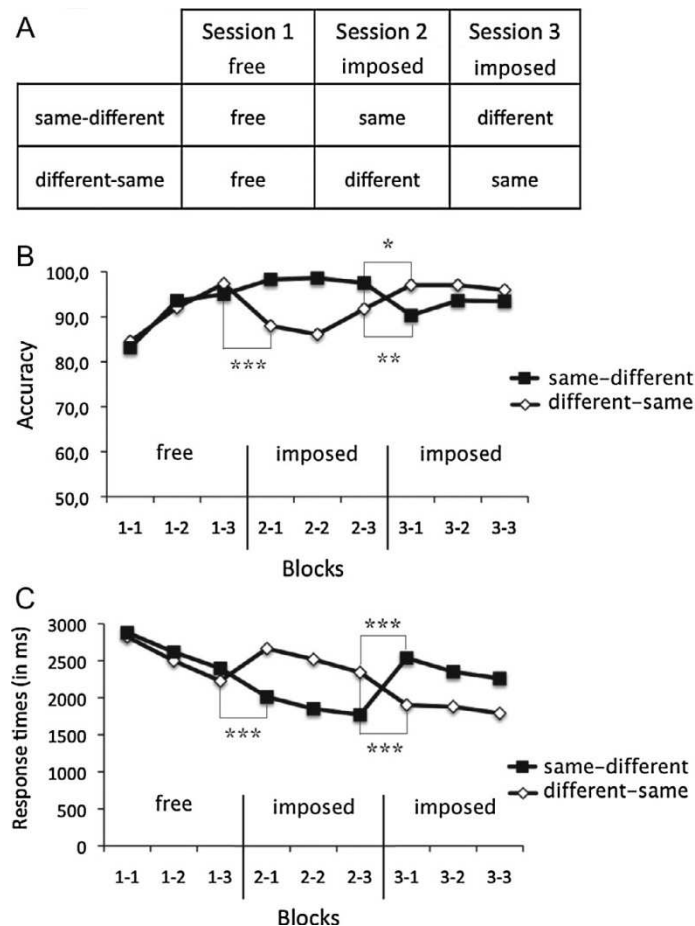


Figure 11. Participants' accuracy (percentage correct) and response times as a function of Block and Order of imposed perspectives (same-different, different-same). For both accuracy and response times, the performance decreased significantly when imposing a different perspective and increased significantly when returning to the same perspective as in Session 1.

To summarize the results of this study, perceivers do not adopt the same perspective when interpreting ambiguous spatial information. Some persons prefer to perceive space from their own centred point of view whereas others prefer the point of view of another person (20% of the cases in our study). In addition, this perspective corresponds to a natural reference frame rather than a mere cognitive choice.

### III.3. The influence of vision on the reference frames that are adopted

In a second study, conducted in collaboration with Gabriel Arnold (post-doctoral researcher), our aim was to investigate the influence of vision on the reference frames that are adopted (Arnold & Auvray, in preparation). To do so, early blind, late blind, and blindfolded sighted participants completed an experiment with same protocol as in Arnold, Spence, & Auvray's 2016 study (described in section III.2), with however some minor adaptations so that it can be completed without sight.

The results on the proportions of responses revealed that blind participants adopted less frequently decentred perspectives than blindfolded sighted control participants ( $\chi^2(2) = 7,33; p = .026$ ). There was no significant difference between early and late blind participants ( $\chi^2(2) = 1,60; p > .448$ ). We then computed the proportions from the three groups of participants in this study, which were labelled as “without vision” condition (early blind, late blind, and blindfolded sighted participants) and we compared them to the results obtained in our previous study, a group which were labelled “with vision”. The participants without vision adopted more frequently an egocentric perspective corresponding to localizing the tactile stimulus on the body. On the other hand, the participants with vision adopted more frequently an egocentric perspective corresponding to a projection of the tactile stimuli out of the body ( $\chi^2(2) = 24,15; p < .001$ ), see figure 12.

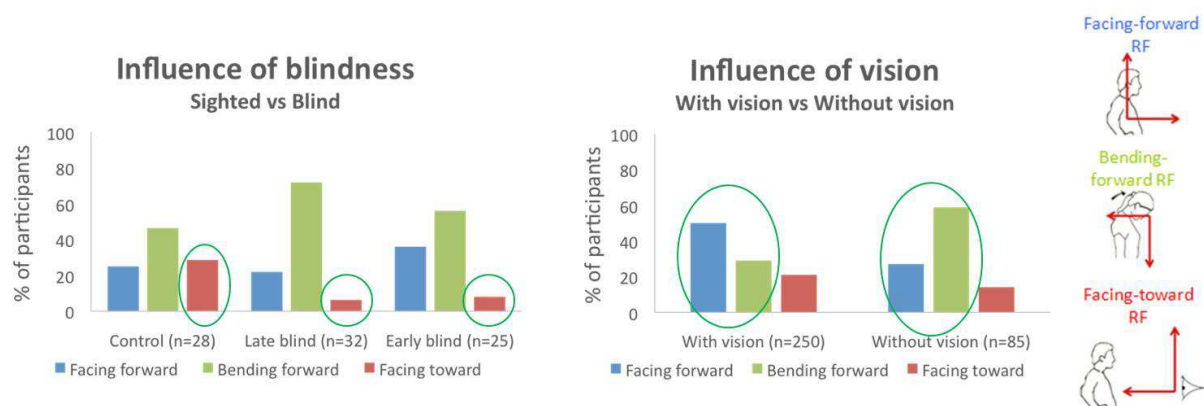


Figure 12. On the left, proportions of the early blind, late blind, and blindfolded sighted participants adopting each of the three perspectives in this study. On the right, proportions of the participants with vision and without vision (early, late blind, and blindfolded sighted) adopting each of the three perspectives.

It should be underlined here that the different self-centred perspectives can correspond to a mental localization of the tactile stimulation either on the body (proximal attribution) or on the external space (distal attribution). Here, the head-centred perspective corresponds to the proximal attribution (the stimuli are considered by the participants to be on the skin) and the trunk centred perspective corresponds to the distal one (the stimuli are considered by the participants to be projected in front of them; see figure 9 for a schematic illustration of the different axes of perception and projections).

We finally computed the cost in accuracy when the participants changed from their preferred perspective to an imposed different one. Early-blind participants showed a greater cost in accuracy when adopting an unnatural perspective than sighted and late-blind participants ( $t(133) = 3,00$ ;  $p = .003$ ), there was no significant difference between sighted and late blind participants ( $t(109) < 1$ ).

To summarize, the results of this study revealed that vision and prior visual experience influence the perspectives that are adopted to interpret tactile stimulation received on the body surface. First, blind people adopt less frequently the perspectives that are not centred on their body (i.e., decentred perspectives). This highlights the important role of vision for spatial cognition and perspective-taking (see Afonso et al., 2014; Pasqualotto et al., 2013, for similar claims with different methodologies). Second, sighted people are more prone to adopt an egocentric perspective which corresponds to projecting the tactile stimulation in front of their body when they can see during the task. However, when they are temporarily deprived of vision by a blindfold, they are biased toward an egocentric perspective which corresponds to localizing the tactile stimulation on the body. The spatial perspectives that are adopted in touch are thus influenced both by the lack of vision, be it from birth or from several years, and by a temporary deprivation of vision.

Note that this study has implications for the design of tactile devices conceived to compensate visual impairments, as the results reflect the importance to take into account individual specificities. These individual specificities are particularly crucial for early-blind people, as they show greater difficulties to adopt an unnatural perspective. One possible solution to overcome this consists in adapting the tactile device to their preferences in orienting left-right and top-bottom of tactile stimulation. In addition, our hypothesis, which is planned to-be-tested, is that blind persons using a visual-to-tactile sensory substitution device might favour a trunk-centred perspective as compared with blind persons not using one, who will remain using preferentially the head-centred perspective.

### III.4. The embodied nature of spatial perspective taking

In the third study, conducted in collaboration with Gabriel Arnold (post-doctoral researcher), our aim was to investigate whether adopting a head-centred perspective on the body is an embodied process (Arnold & Auvray, 2017). The head-centred perspective, which consists of mentally rotating the head toward the stimulated surface, is particularly interesting for investigating embodied perspective-taking as it involves mentally changing the body posture. Crucially, such a head-centred perspective has been described to be constrained by real body movements (Sekiyama, 1991).

In this study, the participants were asked to adopt a head-centred perspective to recognize the letters b, d, p, and q, presented on eight body surfaces (see Figure 13). They were instructed to do not look at the stimulated surface and to keep a standing position with the head oriented forward. They were asked to indicate which letter was recognized. The results of the analyses conducted on accuracy (percentage of correct responses) and response times (RTs) showed that adopting a head-centred perspective on the body is easier for surfaces that can be easily looked at with real head movements. The head-centred perspective was harder to adopt for surfaces that were either far away from the head (e.g., the shin), physically impossible to be directly looked at (e.g., the back), or necessitating a greater quantity of movements (e.g., the sides compared with the front and the legs compared with the trunk).

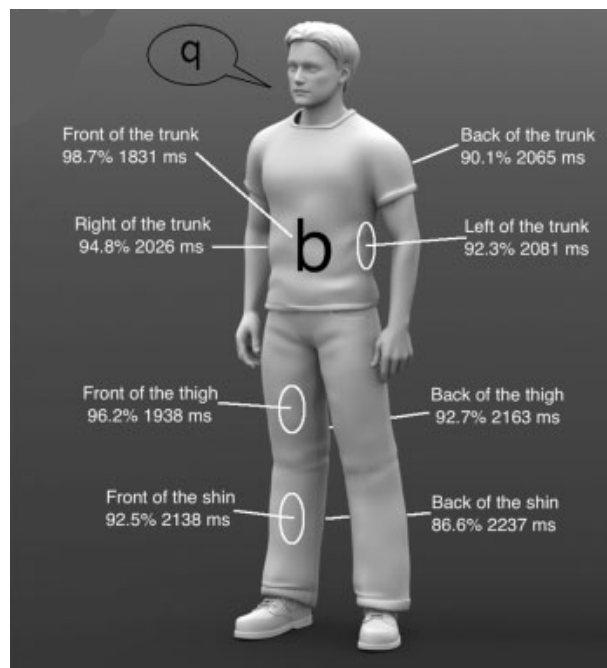


Figure 13. Mean accuracy and RTs for the eight stimulated surfaces.

Overall, the results of this study support the embodied nature of perspective-taking. They show that mentally taking a spatial perspective on our own body involves a mental change in

body posture that is influenced by physical body parameters. More generally, they highlight the important role the body plays in perception.

### **III.5. The role of proprioceptive information on spatial perspective taking**

In the fourth study, conducted in collaboration with Gabriel Arnold, Laura Fernandez (post-doctoral researchers), and Fabrice Sarlegna (UMR 7287) our aim was to investigate the influence of proprioception on the adopted reference frame (Arnold et al., submitted).

Previous studies highlighted that modifications of people's information relative to their body modify the reference frames they adopt. For instance, as was mentioned in section II.1, Ferrè et al.'s (2014) study using the graphesthesia task showed that when receiving galvanic stimulation of the vestibular system, participants adopted more of a self-centred perspective than a decentred one, compared to a control condition. The control condition consisted in galvanic stimulation provided on the neck's skin, which provoked light tactile sensations without vestibular stimulation. The authors suggested that the vestibular system plays an important role in binding together multisensory information into an egocentric body-centred reference frame (see also Lopez, 2016, for a review of the role of vestibular information on spatial cognition). On the other hand, according to the authors, perturbation of the vestibular system may disrupt the unity of the self and the body, giving rise to more decentred perspectives. We made the hypothesis that similarly, proprioceptive deficits might bias toward adopting a decentred perspective.

We used the graphesthesia task (this time with the letters drawn manually) in two deafferented patients (IW and GL) that have neither tactile nor proprioceptive perception below the head and a group of matched control participants that have no perceptual deficits. The ambiguous letters were drawn on the participants' forehead, left, and right surfaces of the head. This was done when their head was oriented forward, leftward, and rightward. We used these surfaces as, when it comes to adopt a spatial perspective, such as deciding whether an object is located to the body's left or right, the global orientation of the self has been shown to correspond to a combination of trunk and head orientations (Alsmith & Longo, 2014). The results obtained in control participants (N=20) revealed that, when the forehead was stimulated, a great majority of responses corresponded to adopting a self-centred perspective. For side surfaces, when the head was oriented forward, there was no preference for adopting a self-centred perspective or a decentred one. However, when the head was oriented leftward or rightward, the side surfaces of the head were aligned with the front and back surfaces of the trunk and there was a significant preference for the self-centred perspective. This confirms that in healthy participants the head and the trunk both play a role in orienting the self. Different results were obtained with deafferented patients, who did not show any effect of the stimulated surface nor of the head orientation on the perspective they took. This suggests that

their physical body does not influence the orientation of their self. Rather, these patients rely on external stimulation and they adopt strategies that are more cognitive and more prone to inter-individual differences. One of the patients consistently adopted a self-centred perspective in all conditions whereas the other consistently adopted a decentred perspective. These results are in line with the inter-individual differences we previously obtained with the same deafferented patients completing a spatial navigation task (Renault, Auvray, Parseihian, Miall, Cole, & Sarlegna, 2018). This study highlights the crucial role of proprioceptive information in defining the location and orientation of the experienced self. Thereby it allows shedding light on the contribution of the different sensory modalities on how the self is experienced.

### III.6. The transfer of a learned reference frame

In this fifth study, conducted in collaboration with Jess Hartcher-O'Brien (post-doctoral researcher), we investigated the perspectives taken on the hand as a function of its orientation, whether the perspective is malleable and can be updated to a novel perspective via adaptation, and finally whether adaptation on one surface transfers to non-adapted surfaces. **Note that this study is included as the sixth of the selected articles in the Annex (Hartcher-O'Brien & Auvray, 2016).**

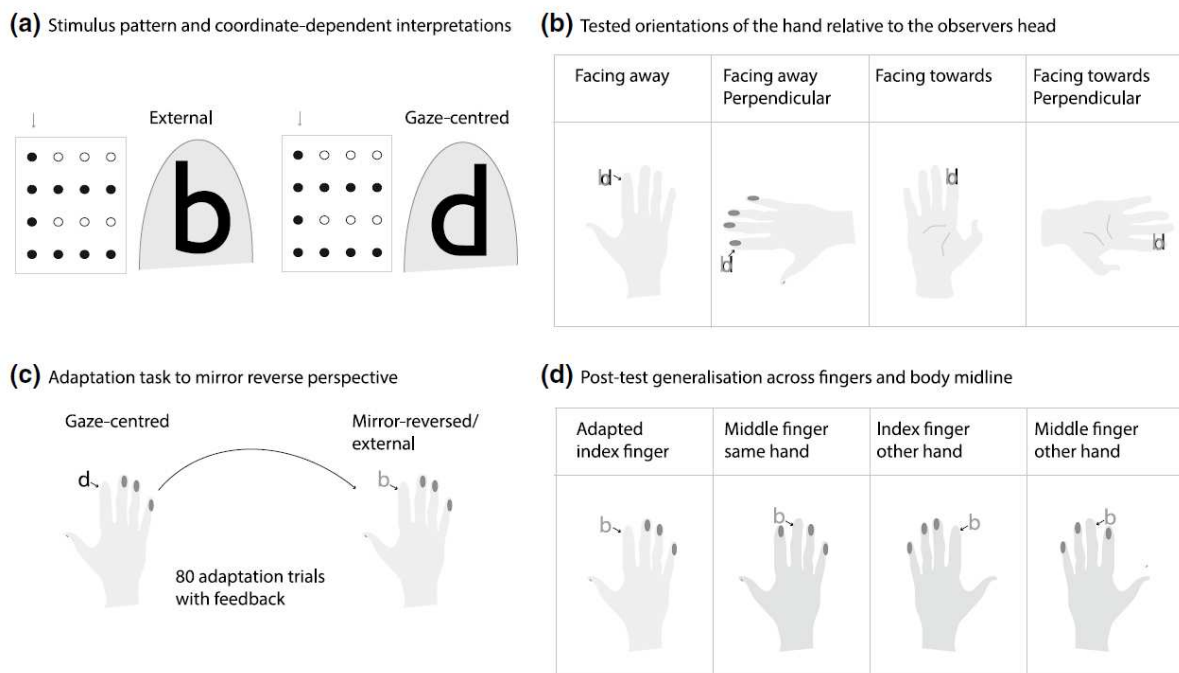


Figure 14. Cutaneous patterns, potential perspectives, and fingertip orientations adopted during the experiment. (a) The device pin activation pattern for the letter *b*. The interpretation of the pattern as a *b* required assigning external coordinates to the activation pattern. An interpretation for a *d* would involve assigning the reverse, gaze-centred, coordinates. (b) The four surface orientations tested during task 1: orientation manipulation. The orientations are given with reference to the head of the

observer. (c) Illustration of the adaptation phase used in task 2. (d) Fingers tested during the post-adaptation phase. Orientation of the fingertip was held constant.

We investigated this specific body surface, because given the ability of the hand to explore objects from different vantage points, independent of the posture of the body, it may well exhibit unique reference frame selection relative to other body surfaces. The participants in our study were presented with the letters b, d, p, and q on their fingertips, by means of a tactile interface (Tactos, see Gapenne et al., 2003). They had to respond which of the letters they recognized. The experiment was divided into 3 tasks. In the first one, the participants were presented with the tactile letters on the index fingertip of their dominant hand and four surface orientations were tested. In the second task, the participants were required to adopt the reference frame corresponding to the device (i.e. mirror reversed as compared to a head-centred perspective) and they were trained with this perspective. In a third task, we measured whether this adapted perspective would generalize across non-adapted fingers and across the body midline (see figure 14).

Our results first showed that the orientation of the hand in the vertical plane determines the perspective taken: an external perspective is adopted when the hand faces the observer and a gaze-centred perspective is selected when the hand faces away. Second, the participants adapt to a mirror-reversed perspective through training. Third, this adapted perspective holds for the adapted surface and generalises to non-adapted surfaces, including across the body midline (see figure 15). This transfer suggests that the process measured here is a higher-level perspective taking, rather than an automatic coordinate assignment process linked to the receptive fields of the stimulated surfaces (Harrar et al., 2014).

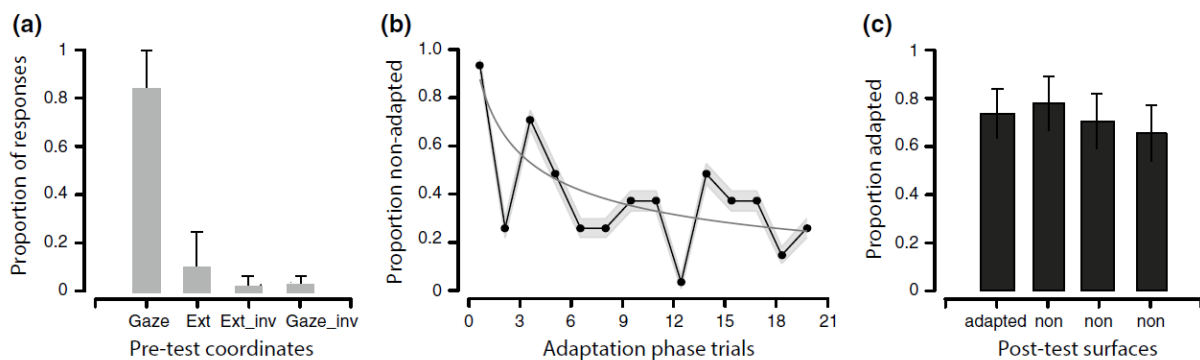


Figure 15. Proportion of responses consistent with the natural and adapted perspectives, before, during, and after the adaptation task. (a) The proportion of responses consistent with the predicted, gaze-centred perspective, and the three other non-predicted ones, averaged across observers. The error bars represent the standard error of the mean. (b) The change in the response pattern during adaptation (averaged across observers and across the four letters). That is, adaptation phase trial 1 corresponds to four trials involving each of the four letters b, d, p, q). The variance (shaded error bar) represents the standard error across observers. The polynomial fit to the proportion of non-adapted responses, represented by the continuous grey line. (c) The proportion of responses in the post-test (averaged across observers and letters) consistent with the adapted perspective for the four surfaces:

the adapted fingertip, the non-adapted middle finger of the adapted hand, the homologous finger of the non-adapted hand, and the middle finger of the non-adapted hand. The error bars represent the standard error of the mean.

Note that the transfer results have implications for training with novel haptic interfaces, suggesting that it is possible to remap, with a limited amount of training, the local cutaneous pattern of stimulation into different coordinates for the purposes of experiencing a three-dimensional world and a novel perspective from your own gaze-/body-centred point of view.

### III.7. The bi-stable palm

A sixth study was conducted in collaboration with Jess Hartcher-O'Brien (post-doctoral researcher) and Vincent Hayward (ISIR, UMR 7722). Its aim was to investigate further the perspectives that are taken on the hand (Hartcher-O'Brien, Hayward, & Auvray, in preparation). This study used a variation of the graphesthesia task. Circular shapes were drawn on the body surface by means of a tactile interface (Tactos, see Gapenne et al., 2003). These stimuli could be drawn clockwise along the device or counter-clockwise. The participants' task was to answer if they perceived each stimulus as being presented clockwise versus counter-clockwise and we identified the reference frame they adopted from their responses. In the first experiment of this study, the stimuli were presented on the index and palm of the participants' dominant hand, which each time varied across 5 orientations (see figure 16).

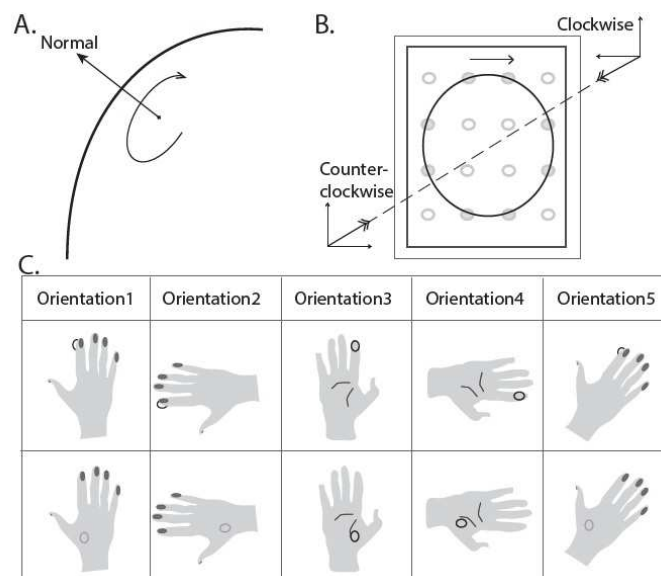


Figure 16. Viewpoint selection, stimuli and procedure. (A) The clockwise/counter-clockwise pattern as displayed on the surface of the device. (B) Experiment 1. Observers ( $n = 16$ ) placed their dominant hand index finger or palm against the active surface of the device and reported whether the pattern moved clockwise or counter-clockwise in a 2AFC procedure. (C) The five orientations tested.



The obtained results revealed that when the stimulation is received on the index, the reference frame that is adopted varies with the orientation of the hand. As in Hartcher-O'Brien & Auvray (2016), an external perspective is adopted when the hand faces the observer and a gaze-centred perspective is selected when the hand faces away. A less straightforward response pattern was observed for stimulation on the palm, suggesting a potential change in processing across the skin's surfaces for orientation (see figure 17). This finding is in contrast to what has previously been assumed, i.e. that perspective taking across the hand is homogeneous (Parsons & Shimojo, 1989).

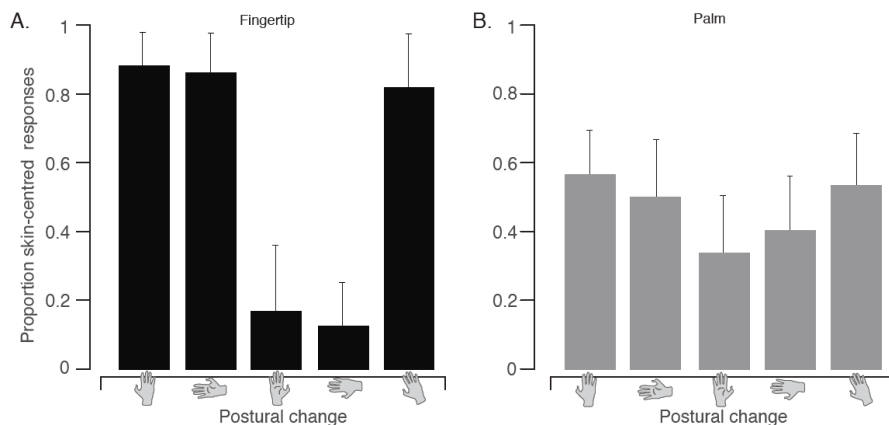


Figure 17. Viewpoint results for hand orientation. Proportion of gaze-centred responses for the five orientations tested for the fingertip (A) and the palm (B) averaged across participants. The error bars represent the standard deviation.

Our aim in the next experiment was to verify if this response pattern identified for the palm could potentially be the result of poorer sensory acuity, i.e. lower resolution within this neighbourhood than for the fingertip (Mancini et al., 2014; Stevens & Choo, 1996) or if, instead it reflects a different process. To do so, the fingertip and then the palm were again stimulated with a circular pattern of motion, this time containing gaps in one of four possible locations. The participants' task of was to identify the position of the gap in the pattern of stimulation. Only clockwise patterns were used in this experiment and a control, 'no gap' stimulus allowed us to calculate  $d'$ prime. The results revealed that although the sensitivity was higher on average on the palm than on the fingertip, such sensitivity could not predict the coordinate assignment pattern on the fingertip nor on the palm ( $p > .05$ , see figure 18). This result suggests that the difference across skin surfaces cannot be accounted for by differences in spatial acuity of the two regions but is more likely linked to the functional role of the surface in everyday interactions.

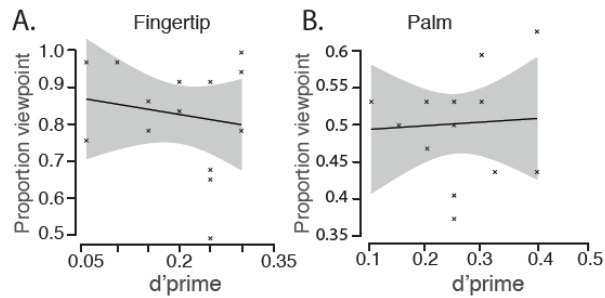


Figure 18. Correlation between change in viewpoint responses and the sensitivity of each surface as measured via the gap detection procedure for the fingertip (A) and the palm (B). The correlation between sensitivity,  $d'$ prime, and proportion change in perspective is not significant for any of the surfaces tested.

The third experiment investigated whether viewpoint selection across the entire surface of the hand was continuous or discretely separable into functionally defined neighbourhoods. The same procedure as in Experiment 1 was used, this time on eight neighbouring locations on the volar surface of the hand. If the response pattern changes gradually from fingertip to palm this would suggest continuous, non-functional organisation of the skin boundary. If, on the other hand, the response pattern changes discretely across regions, from finger to palm, this would support a functional organisation of the surface. Figure 19 shows the discrete changes from the finger region to that of the palm. With respect to viewpoint adopted, three clusters emerged. The three groups are distinguishable and suggest a function-based viewpoint selection process on the hand.

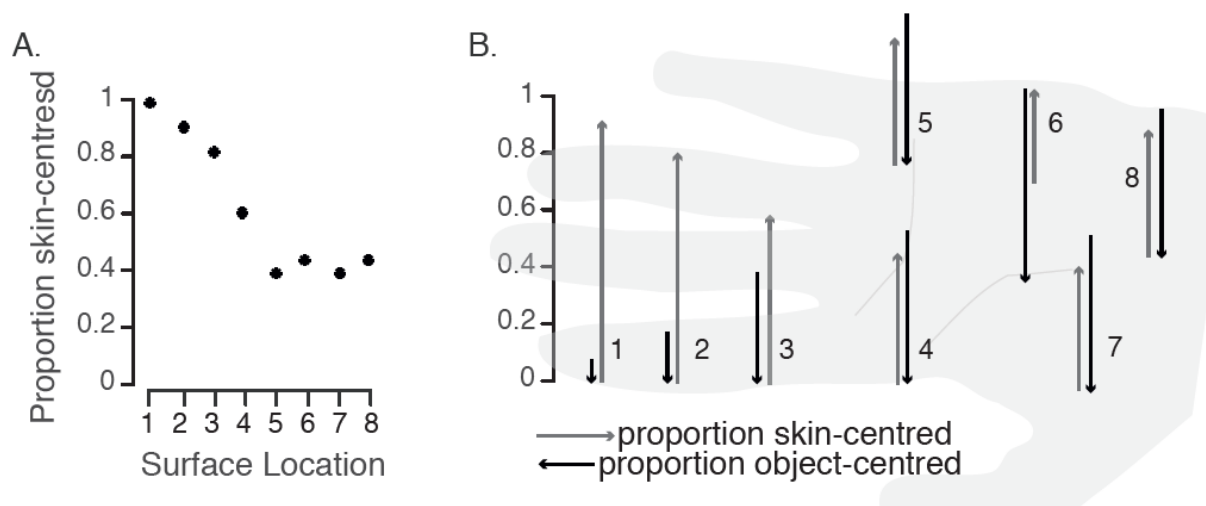


Figure 19. Role of surface function via structural regions of the hand. The clockwise-counter clockwise stimuli were applied to distinct regions systematically across the hand. The proportions of body centred responses are given as a function of the stimulated locations.

### III.8. The processes of distal attribution

One of the key mechanisms of spatial cognition is distal attribution, also known as exteriorization. When tactile stimuli are used, distal attribution corresponds to the processes

allowing attributing the stimuli received on the body surface to an object located in the three-dimensional space. The research described in this section was conducted in collaboration with Vincent Hayward (ISIR, UMR 7722), Jessica Hartcher O'Brien, and Alexander Terekhov O'Brien (post-doctoral researchers).

It first should be mentioned that distal attribution is of important interest in sensory substitution. Indeed, one interesting result obtained with these devices suggests a change in users' subjective experience. After training, they no longer report feeling the stimulation on their skin, where it occurs, but directly attribute the cause of the stimulation to a distant object (Bach-y-Rita et al., 1969; Bach-y-Rita and Kerckel, 2003; White et al., 1970;). **Note that a review on distal attribution in sensory substitution is included as the seventh of the selected articles in the Annex (Hartcher-O'Brien & Auvray, 2014).**

In addition, it should be mentioned that the graphesthesia task was one of the first paradigms we thought of to investigate distal attribution. What can be said for now about the graphesthesia task is that adopting a head-centred perspective for interpreting tactile stimuli does not necessarily imply the conversion of the stimuli into a visual format. However, both visual and tactile perspective taking seem to involve a common externalization process: the origin of the perspective must be located at a different position than the object of perception. In vision, the object is external to the perceiver, located in front of the head. In touch, the externalization process involves projecting the tactile stimulation forward, centring the perspective on another body part than the one being stimulated (i.e., bending the head forward to see the stimulation), or adopting a decentred perspective. The adoption of a consistent spatial perspective on tactile stimulation may thus characterize the transition from experiencing the tactile stimulation on the skin (proximal attribution) to becoming able to gain access to the distant object represented by the tactile stimulation (distal attribution, see Arnold, Spence, & Auvray, 2017). This distal attribution process, also named referral of touch (e.g., Petkova & Ehrsson, 2009), is crucial for the ability to distinguish internal from external stimulation, and more generally, the self from the external world. We plan to investigate in the coming year changes in this process with the graphesthesia task, before and after training with a visual-to-tactile sensory substitution device.

A study was also conducted on haptic shape constancy across distance, in collaboration with Jessica Hartcher O'Brien, Alexander Terekhov (post-doctoral researchers), and Vincent Hayward (ISIR, UMR 7722, see O'Brien, Terekhov, Auvray, & Hayward, 2014). Perceptual constancy is central to the process of distal attribution as it allows the nervous system to maintain a coherent experience of the world despite being able to move its sensors independently to the environment. In vision, constancy refers to the stability of the perceived properties of an object over space and time which can be achieved by converting retinal

signals into spatiotopic coordinates, allowing the observers to perceive space independently of their eye movements (Holway & Boring, 1941; Wexler & Held, 2005). In haptic perception, the problem is tricky as well, given that the sensory surfaces undergo even more complex movements in space. To give examples, an object's perceived haptic size is influenced by the extension of the arm, as well as by local deformation of the probe (e.g., Dostmohamed & Hayward, 2005). Thus, any sensory system with the ability to move and explore the environment faces the problem of accounting for changes in proximal patterns of stimulation that are due to self-generated movement, and those that are not. Moreover, in order to experience objects in external space, these proximal patterns need to be converted into an external / spatiotopic frame of reference such that the object can be represented independently of the observer generated perturbations.

The aim of this study was to investigate whether perceived curvature changes as a function of object displacement from the observer. To do so, blindfolded sighted participants had to scan consecutively the surface of two shapes and to report which of the two objects had higher curvature. One shape was close to the observer and the other was positioned further away. This spatial displacement changes the available proprioceptive information about the object shape, and therefore the combined proprio-tactile information may signal different objects at the two distances. The stimuli consisted in 13 three-dimensional rectangular forms with a circular hole. The curvature of one stimulus, the standard, was always maintained constant and the curvature of the others varied in steps of  $2 \text{ m}^{-1}$  (see figure 20).

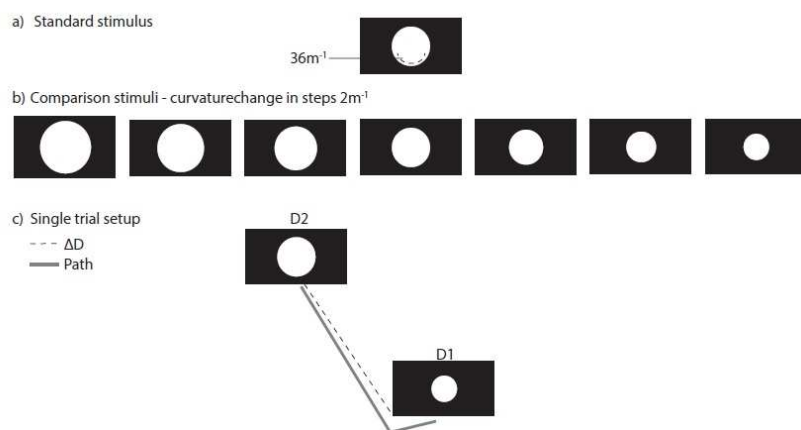


Figure 20. Schematic representation of a) the standard stimulus and b) a subset of the comparison stimuli with changes in curvature by a step of  $2 \text{ m}^{-1}$ . c) Single trial setup for the 2AFC procedure in which participants compared the curvature of two objects: one located at D1 and the second positioned at D2. The exploration was done consecutively with one hand, observers followed a line of tape on the tabletop to guide them to the second object.

The results reveal a perceptual compensation for the change in proprioceptive information. However, two distinct patterns of distance compensation emerged. For one group, haptic shape estimates across distance were consistent with visual-like compensation, in other words, closer shapes needed to have lower curvature to be perceived as the same shape as

the standard at D2. The other group showed a reversed pattern of response such that objects further away need to have lower curvature to be perceived having equal curvature. In this case, the participants performed a reweighting of the proprio-tactile information to account for the displacement (see figure 21).

These results suggest that perceived haptic curvature across distance depends on observers' differential weighting of the multiple available cues. Therefore unlike in visual processing there are multiple sources of information that can weigh in to account for the displacement. One hypothesis for the difference observed is that one group relies more on local touch information, while the second group weights proprioceptive cues more highly.

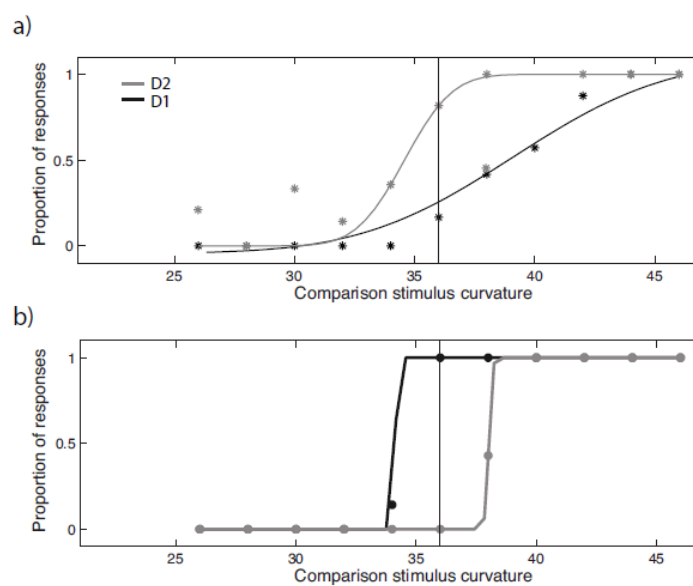


Figure 21. Psychometric functions for estimated curvature as a function of distance: when the standard is close at D1 (grey) and then further from the observer at D2 (black). a) An example observer from group 1 for whom the function for the close object shifts towards high curvature while distant objects are more likely to be perceived as low curvature. b) An example observer from the subset showing the reverse pattern for near and far objects. Data points in the fitting procedure have been weighted according to the number of trials used to obtain them.

### III.9. Conclusions on multisensory reference frames

Katz (1925) suggested that touch has many of the capabilities of a distal or ‘far’ sense, such as vision and audition. However, spatial properties of objects, such as their distance from the observer, are readily available through the ‘far senses’ but not to the ‘proximal’ ones and in this way our experience of space depends on the sensory receptors stimulated. For tactile perception, attribution to a distal source or to the body surface can differ depending on the circumstances. For instance, active exploration via touch tends to promote distal attribution (Gibson, 1962; Katz, 1925). Characterizing spatial perception in touch thus requires to understand how the proximal stimulation (i.e., the stimulation received at the receptor

surface) spatially relates to its cause. In the experiments described in this section, we first approached this question by using the graphesthesia task to interrogate the coordinate systems that are used to perceive tactile stimuli. To start (III.2), our results revealed that there is an individual variability with some persons preferring to perceive space from their own centred point of view and others preferring a decentred perspective (20% of the cases in our study). This reference frame is natural and does not reduce to a mere cognitive choice. Second (III.3), our studies revealed an influence of vision on the reference frames that are adopted with blind people adopting less frequently the decentred perspectives and the (temporary or permanent) lack of vision favouring the egocentric perspective. (III.4) We also showed that the head-centred perspective has an embodied nature, as it is easier to adopt for surfaces that can be easily looked at with real head movements. (III.5) In addition, deafferented patients rely on external stimulation and they adopt strategies that are more cognitive and more prone to inter-individual differences. Regarding the hand, our studies showed (III.6) that when the participants are trained to adopt a novel reference frame on one body surface, there is a transfer of learning to other body surfaces. In addition, an external perspective is adopted when the hand faces the observer and a gaze-centred perspective is selected when the hand faces away. Far more variability was found when stimuli are presented on the palms than on the fingertips. (III.7) This variability in the responses was found to change discreetly across regions, from finger to palm, which supports a functional organisation of the surface. Finally (III-8), using a different paradigm of curvature estimates, we found an inter-individual variability with some persons relying more on local touch information and others more on proprioceptive information.

These studies have implications for the design of tactile devices conceived to compensate for visual impairments, as they reflect the importance to take into account individual specificities. The use of these devices can be eased by implementing flexible options as a function of the differences in the reference frames adopted by people (egocentric versus allocentric; trunk-centred versus head-centred). These individual specificities are particularly crucial for early-blind people, as they show greater difficulties to adopt unnatural perspective. One possible solution to overcome this consists in adapting the tactile device to their preferences in orienting left-right and top-bottom of tactile stimulation. The transfer results have implications for training with novel haptic interfaces, suggesting that it is possible to remap, with a limited amount of training, the local cutaneous pattern of stimulation into different coordinates for the purposes of experiencing a three-dimensional world and a novel perspective from a gaze-/body-centred point of view.

The graphesthesia paradigm can also have implications for understanding the social and personal factors that influence the perspective that people adopt on their own body. This aspect will be detailed in the section perspective of research. Finally, further research with the

graphesthesia task and other paradigms investigating distal attribution might allow shading light on the distinction between sensation and perception. Indeed, distal attribution occurs when we experience a pattern of stimulation and assign the cause of this stimulation to an object, stimulus, or event located in external three-dimensional space. When proximal stimulation is not assigned an external cause, it gives rise to a mere sensation, not the perception of an object located at a distance. Therefore understanding how distal attribution emerges is central to understanding the distinction between bodily and external space, as well as the plasticity of this border.

#### **IV. Perspectives for future research**

The scientific research I plan to conduct the coming years aim at investigating the plasticity of our perceptual and bodily representations during tool use (involving the specific tools that are sensory substitution devices) and the role of multisensory information on this plasticity. This research schedule will be presented divided into three themes: 1) Exteriorisation and embodiment; 2) Sensory substitution of vision and proprioception; 3) Interactions between spatial and mental perspectives.

##### **IV.1. Exteriorisation and embodiment**

Bodily and distant spaces seem well defined and delineated. There is the space of my body, the place where I feel, where I receive sensations, which is marked by the boundaries of my skin and the external space made up of objects that I perceive as being out there. However, several situations reveal that such distinction is not straightforward. For instance, when using a virtual reality or a tele-presence device I feel as being at a different place than where my body actually is. These examples underline the possibility to modify the boundaries of the body and hence its plasticity. In order to understand in full such plasticity, two sides of the same question have to be considered jointly: the attribution of distant stimuli to the body space and the attribution of proximal stimuli to a distant space. The research I have conducted so far focused on the second side of the question. My aim is to both prolong this line of research and to investigate its counterpart, the phenomena of bodily extensions and modifications.

Regarding the mechanisms of distal attribution or exteriorisation, we are currently extending the investigation of how tactile stimuli received on the body surface are projected to a distant space to the study of the effect of the spatial coincidence between eye fixations and tactile stimuli. This work started in July 2018 thanks to a post-doctoral funding from Oculus research (PI: V. Hayward). If the results are encouraging, we will then extend this work to the broader study of tactile sensation projections in link with oculomotor signals.

Regarding the investigation of bodily extensions and modifications, first, in collaboration with Vincent Hayward (ISIR) and Gabriel Arnold (company Caylar, previously post-doctoral researcher at ISIR), we started to investigate the numbness illusion (e.g., see Dieguez et al., 2009, see also figure 22). We believe that the effects of this illusion are as promising as those obtained with the rubber hand illusion while allowing for a diversity of new objective measures (such as the use of accelerometers, the measure of precision grip). A first investigation, during the internship of the master student Nathan Rish (2017-2018), interestingly showed that people touch themselves differently before and after the illusion. This promising line of research also validated the use of accelerometers (to measure forces of pressure) to investigate this question. Many control conditions remain to be conducted and, once this is done, we plan to extend this line of work to the shared factors that underlie the illusion as well as the individual differences at stake. The shared factors can include for instance hand configuration (i.e., if my hand is considered to be in my space or in the other's one, see figure 22). Regarding the individual factors, the illusion can induce two different feelings; some persons feel that their finger becomes numb, as if they were giving it to the partner whereas other persons feel an extension of their finger as if they incorporated that of the partner. The question then arises as to what account for these individual differences (e.g., in terms of spatial abilities, interpersonal skills, cognitive style). More broadly, our aim with this illusion is to investigate the link between body perception and body schema, through questions such as: If the body image is modified by the numbness illusion, does it change the way we touch ourselves? Does it change the way we touch someone else? If our body image determines our body schema, is there a difference between touching ourselves and touching someone else?



Figure 22. Illustration of the numbness illusion. The participant rubs his index finger and that of the experimenter (around 10 times). The illusion consists in feeling either a larger or a numb finger. Here are presented two possible configurations: the classic one and the reverse one.

The mechanisms underlying body plasticity will also be conducted within the project “developmental tool mastery” (funded by the ANR, PI: Alessandro Farnè). The project aims at offering a comprehensive multidisciplinary framework for unveiling the rules of body representations plasticity that allow humans to master tool-use. More precisely, we aim at



investigating how young humans acquire control of tools to reach and grasp objects and to sense and explore their environment. Then we will turn to the question of how the adult human brain does code for these abilities. Finally, we will address the question of how do multiple sensory modalities contribute to acquiring and regulating these abilities. In co-supervision with Vincent Hayward, we have hired a post-doctoral researcher for two years in November 20018 at ISIR to work on these question. We have hired a second post-doctoral researcher for the same period to work on closely related topics. In particular, the aim will be to use novel techniques (virtual and tactile interfaces) to empirically investigate how tool-use sensing is encoded in the brain, by disentangling the sources of information arising from physical invariants. This will be done by creating paradoxical stimuli that clearly violate these invariants, and by selectively suppressing sources of information or sensory cues arising from one modality (haptics, vision, audition) while maintaining and even enhancing the other sources.

#### **IV.2. Sensory substitution of vision and proprioception**

The second theme of the research planned prolongs my previous work on sensory substitution devices and extends it to devices designed to compensate for proprioceptive impairments.

##### Sensory substitution of vision

First, I plan to carry on the study of the crossmodal correspondences in sighted and blind persons (described in section II-4). The goal is now to investigate more systematically the possible crossmodal correspondences in audition, vision, and touch, focusing on the dimensions of the signals that are relevant for sensory substitution, and in addition, to investigate the ways in which these correspondences are combined together.

Second, an intrinsic problem with systems encoding visual information into a single sensory modality is that a maximum of 3 dimensions of the visual signal can be encoded simultaneously. Hence, whereas some devices only code for depth (z-axis), others code for the horizontal and vertical axes (x & y axes) together with brightness. None of them can provide a code for all three axes and brightness simultaneously. This can be achieved by developing new multisensory substitution devices that use the input of more than one sensory channel. In collaboration with Maurice Ptito (McGill, Montreal), Ron Kupperts (University of Copenhagen), Ophelia Deroy (University of Munich), and the company Roger McMahon & Associates (Glasgow), our aim is to develop the I-VEST (Innovation through Vision Enhancement by Sound and Touch), a multisensory augmentation device consisting of a tight fitting shirt with integrated tactile stimulators (actuators) and a new generation of headphones. A mini-camera, mounted on a pair of glasses will capture visual information that will be converted into tactile and auditory code, and wirelessly transmitted

to the ears and skin of the upper torso. Thus, here, visual information will be translated into both tactile and auditory code which will allow us to convert all three spatial axes and brightness. To do so in an optimal way, we will first investigate the audio-tactile interactions in blind and sighted individuals, to measure the neurobehavioral impact of the absence of vision on these interactions. Then we will integrate our findings into our portable and integrated technologies based upon audio-tactile information processing. Our aim is to build this device for blind individuals, for those suffering from other visual impairments, and for sighted people in conditions of low visibility. One key aspect is also to borrow methodologies from the video game and the entertainment sector to develop novel approaches to sensory augmentation. We plan to apply to European grants to fund this project.

#### Sensory substitution of proprioception

Two projects are starting regarding the compensation for proprioception. The first one, I-Gait (PI: Wael Batcha, ISIR, funded by ANR) aims at better preventing falls in both elderly people and people suffering from proprioceptive deficits by means of a device assisting them when they walk. We will develop a tactile handle, which delivers computer-generated tactile feedback able to initiate the necessary postural and force adjustments, mounted on the top of a robotic cane. Our aim is to investigate the best conversion of proprioceptive signals into tactile ones such that users can initiate automatic and precise adjustments of the postural balance and of the force applied to the cane by mechanically stimulating the skin of the hand palm. This will be built on previous work (Terekhov & Hayward, 2015; V  rit   et al., 2014) which showed that people's centre of pressure can be displaced precisely, without any conscious cooperation, through an appropriate control of the motion of a surface, which users touch lightly with their forefinger

The second project is conducted in collaboration with Nathana  l Jarasse (ISIR) and Gabriel Arnold (company Caylar). Our aim is to ease the learning of prostheses involving myoelectric control in upper limb amputees. To do so, our aim is to investigate the best way to code for proprioceptive cues into tactile ones that will be transmitted to users by means of a vibrotactile bracelet. We have co-supervised a second year master student (Yann Kechabia) in 2017-2018 to study the coding of wrist rotation into vibrations. The results are promising and our aim now is to extend this work to the other proprioceptive cues that are involved in the use of the prosthesis.

#### **IV.3. Interactions between spatial and mental perspectives**

The aim of this theme is to prolong the research on multisensory spatial perspectives (described in section II) in order to investigate the reciprocal influence of mental and spatial perspectives.

Spatial cognition is subject to important inter-individual variability (e.g., Berthoz & Viaud-Delmon). It is thus important to understand the personal factors underlying this difference. First, during the internship of a master student in clinical psychology (Sandra Inard, in 2016-2017) and in collaboration with Gabriel Arnold (company Caylar), we used the graphesthesia task in participants and correlated their results to those they obtained in personality and attachment questionnaires. Promising results started to emerge, in particular in the link between spatial perspectives, introversion, and attachment. Our aim is now to investigate these links by means of more fine grained questionnaires.

Second, the ability to adopt someone else's perspective develops in infancy. However, this ability is impaired in some persons, such as those in the autistic spectrum. In collaboration with David Cohen and Jean Xavier (MD, G.H.U. Pitié-Salpêtrière) our aim is to use the graphesthesia task in autistic and neurotypical children (aged between 3 and 18 years old). This will allow us to investigate more in depth the link between mental and spatial perspectives. The next step will be to investigate if the reciprocal influence of spatial and mental perspectives can be modulated. In other words, the aim will be to determine the extent to which it is possible to change of mental perspectives by changing spatial perspectives and vice versa.

## V. References

- Abboud, S., Hanassy, S., Levy-Tzedek, S., Maidenbaum, S., & Amedi, A. (2014). EyeMusic: Introducing a “visual” colorful experience for the blind using auditory sensory substitution. *Restorative Neurology and Neuroscience*, *32*, 247-257.
- Afonso, A., Blum, A., Katz, B. F. G., Tarroux, P., Borst, G., & Denis, M. (2010). Structural properties of spatial representations in blind people: Scanning images constructed from haptic exploration or from locomotion in a 3-D audio virtual environment. *Memory and Cognition*, *38*, 591-604.
- Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Science*, *8*, 457-464.
- Alsmith, A. J., & Longo, M. R. (2014). Where exactly am I? Self-location judgements distribute between head and torso. *Consciousness and Cognition*, *24*, 70-74.
- Arno, P., De Volder, A. G., Vanlierde, A., Wanet-Defalque, M. C., Streel, E., Robert, A., et al. (2001). Occipital activation by pattern recognition in the early blind using auditory substitution for vision. *Neuroimage* *13*, 632-645.
- Arnold, G., & Auvray, M. (2018/to appear). Tactile recognition of visual stimuli: Specificity versus generalization of perceptual learning. *Vision Research*.
- Arnold, G., & Auvray, M. (2014). Perceptual learning: Tactile letter recognition transfers across body surfaces. *Multisensory Research*, *27*, 71-90.
- Arnold, G. & Auvray, M. (in preparation). Blind and sighted people differ in the way they mentally represent space from different perspectives.

- Arnold, G., Sarlegna, F., Fernandez, L., & Auvray, M. (submitted). The physical body plays a different role in locating and orienting the experienced self in deafferented patients and in controls.
- Arnold, G., Spence, C., & Auvray, M. (2016). Taking someone else's spatial perspective: Natural stance or effortful decentering? *Cognition*, *148*, 27-33.
- Arnold, G. & Auvray, M. (2017). The graphesthesia paradigm: drawing letters on the body to investigate the embodied nature of spatial perspective taking. *I-Perception*, *8*, 1-5.
- Arnold, G., Pesnot-Lerousseau, J. & Auvray, M. (2017). Individual differences in sensory substitution. *Multisensory Research*, *6*, 579-600.
- Arnold, G., Spence, C., & Auvray, M. (2017). A unity of the self or a multiplicity of locations? How the graphesthesia task sheds light on the role of spatial perspectives in bodily self-consciousness. *Consciousness & Cognition*, *56*, 100-114.
- Aspell, J. E., Heydrich, L., Marillier, G., Lavanchy, T., Herbelin, B., & Blanke, O. (2013). Turning body and self inside out: Visualized heartbeats alter bodily self-consciousness and tactile perception. *Psychological Science*, *24*, 2445-2453.
- Auvray, M., & Farina, M. (2017). Patrolling the boundaries of synaesthesia: A critical appraisal of transient and artificially-acquired forms of synaesthetic experiences. In O. Deroy (Ed.). *Synaesthesia: Philosophical & Psychological Challenges* (pp. 248-274), Oxford: Oxford University Press.
- Auvray, M., Hanneton, S., & O'Regan, J. K. (2007). Learning to perceive with a visuo-auditory substitution system: Localization and object recognition with The Voice. *Perception*, *36*, 416-430.
- Auvray, M., & Myin, E. (2009). Perception with compensatory devices. From sensory substitution to sensorimotor extension. *Cognitive Science*, *33*, 1036-1058.
- Bach-y-Rita, P., Collins, C. C., Saunders, F. A., White, B., & Scadden, L. (1969). Vision substitution by tactile image projection. *Nature*, *221*, 963-964.
- Bach-y-Rita, P., & Kercel, S. W. (2003). Sensory substitution and the human-machine interface. *Trends in Cognitive Sciences*, *17*, 541-546.
- Bach-y-Rita, P., Tyler, M. E., & Kaczmarek, K. A. (2003). Seeing with the brain. *International Journal of Human-Computer Interaction*, *2*, 285-295.
- Berthoz, A., & Viaud-Delmon, I. (1999). Multisensory integration in spatial orientation. *Current Opinion in Neurobiology*, *9*, 708-712.
- Blackmore, S. (1987). Where am I? Perspectives in imagery and the out-of-body experience. *Journal of Mental Imagery*, *11*, 53-56.
- Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience*, *13*, 556-571.
- Blanke, O., Ortigue, S., Landis, T., & Seeck, M. (2002). Neuropsychology: Stimulating illusory own-body perceptions. *Nature*, *419*, 269-270.
- Block, N. (2003). Tactile sensation via spatial perception. *Trends in Cognitive Sciences*, *7*, 285-286.
- Bologna, G., Deville, B., & Pun, T. (2009). On the use of the auditory pathway to represent image scenes in real-time. *Neurocomputing*, *72*, 839-849.

- Boussaoud, D., & Bremmer, F. (1999). Gaze effects in the cerebral cortex: Reference frames for space coding and action. *Experimental Brain Research*, 128, 170-180.
- Boyer, E. O., Babayan, B. M., Bevilacqua, F., Noisternig, M., Warusfel, O., Roby-Brami, A., Hanne-ton, S., & Viaud-Delmon, I. (2013). From ear to hand: the role of the auditory-motor loop in pointing to an auditory source. *Frontiers in Computational Neurosciences*, 7, 26.
- Brunet, L. Darses, F., & Auvray, M. (2018). Strategies and needs of blind pedestrians during urban navigation. *Le Travail Humain*, 81, 141-171.
- Clark, A. (2003). *Natural-born cyborgs: Minds, technologies, and the future of human intelligence*. New York: Oxford University Press.
- Cohen, Y. E., & Andersen, R. A. (2002). A common reference frame for movement plans in the posterior parietal cortex. *Nature Reviews Neuroscience*, 3, 553-562.
- Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: evidence for dual-route models of reading. *Journal of Experimental Psychology Human Perception and Performance*, 20, 1197-1211.
- Danilov, Y., Tyler, M., Skinner, K., Hogle, R., & Bach-y-Rita, P. (2007). Efficacy of electro-tactile vestibular substitution in patients with peripheral and central vestibular loss, *Journal of Vestibular Research*, 17, 119-130.
- Denis, M. (2018). *Space and spatial cognition: A multidisciplinary perspective*. Abingdon, UK: Routledge.
- Deroualle, D., Borel, L., Devèze, A., & Lopez, C. (2015). Changing perspective: The role of vestibular signals, *Neuropsychologia*, 79, 175-185.
- Deroy, O., & Auvray, M. (2012). Reading the world through sensory substitution devices. *Frontiers in Theoretical and Philosophical Psychology*, 3, 457.
- Deroy, O., & Auvray, M. (2014). A crossmodal perspective on sensory substitution. In S. Biggs, M. Matthen, & D. Stokes (Eds.), *Perception and Its Modalities* (pp. 327-349), Oxford: Oxford University Press.
- Deroy, O., Fasiello, I., Hayward, V., & Auvray, M. (2016). Differentiated audio-tactile correspondences in sighted and blind individuals, *Journal of Experimental Psychology: Human Perception and Performance*, 42, 1204-1214.
- Deroy, O., & Spence, C. (2013). Why we are not all synesthetes (not even weakly so). *Psychonomic Bulletin & Review*, 20, 643-664.
- Dieguez, S., Mercier, M. R., Newby, N., & Blanke, O. (2009). Feeling numbness for someone else's finger. *Current Biology*, 19, R1108-R1109.
- Diot, B., Halavackova, P., Demongeot, J., & Vuillerme, N. (2014). Sensory substitution for balance control using a vestibular-to-tactile device, *Multisensory Research*, 27, 313-336.
- Dostmohamed, H., & Hayward, V (2005). Trajectory of contact region on the fingerpad gives the illusion of haptic shape. *Experimental Brain Research*, 164, 387-394.
- Farcy, R., Leroux, R., Jucha, A., Damaschini, R., Grégoire, C., & Zogaghi, A. (2006). Electronic travel aids and electronic orientation aids for blind people: Technical, rehabilitation and everyday life points of view. In M. A. Hersh (Ed.), *Proceedings of the Conference & Workshop on Assistive Technology for Vision and Hearing Impairment*.

- Ferrè, E. R., Lopez, C., & Haggard, P. (2014). Anchoring the self to the body: Vestibular contribution to the sense of self. *Psychological Science*, *25*, 2106-2108.
- Gallay, M., Denis, M., & Auvray, M. (2013). Navigation assistance for blind pedestrians: Guidelines for the design of devices and implications for spatial cognition. In T. Tenbrink, J. Wiener, & C. Claramunt (Eds.), *Representing space in cognition: Interrelations of behaviour, language, and formal models*, Oxford: Oxford University Press.
- Gapenne, O., Rovira, K., Ali Ammar, A., & Lenay C (2003). Tactos: special computer interface for the reading and writing of 2D forms in blind people. *Universal Access in HCI, Inclusive Design in the Information Society*, *10*, 1270-1274
- Gibson, J. J. (1962). Observations on active touch, *Psychological Review*, *69*, 477-490.
- Hanneton, S., Auvray, M., & Durette, B. (2010). The Vibe: A versatile vision-to-audition sensory substitution device. *Applied Bionics and Biomechanics*, *7*, 269-276.
- Harrar, V., & Harris, L. R. (2010). Touch used to guide action is partially coded in a visual reference frame. *Experimental Brain Research*, *203*, 615-620.
- Harrar, V., Spence, C., Makin, T. (2014) Topographic generalization of tactile perceptual learning. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 15-23.
- Heed, T., Buchholz, V. N., Engel, A. K., & Röder, B. (2015). Tactile remapping: from coordinate transformation to integration in sensorimotor processing. *Trends in Cognitive Sciences*, *19*, 251-258.
- Ho, C., & Spence, C. (2007). Head orientation biases tactile localization. *Brain Research*, *1144*, 136-141.
- Ho, C. & Spence, C. (2008) *The Multisensory Driver: Implications for ergonomic car interface design*. Aldershot: Ashgate Publishing.
- Holway, A.H., & Boring, E.G. (1941). Determinants of apparent visual size with distance variant. *American Journal of Psychology*, *54*, 21-37.
- Humphrey, N. (1992). *A History of the Mind*. New York: Simon & Schuster.
- Hurley, S., & Noë, A. (2003). Neural plasticity and consciousness. *Biology and Philosophy*, *18*, 131-168.
- James, W. (1890). *Principles of Psychology*. Holt, New York.
- Jansson, G. (1983). Tactile guidance of movement. *International Journal of Neurosciences*, *19*, 37-46.
- Jeter, P. E., Doshier, B. A., Liu, S. H., & Lu, Z. L. (2010). Specificity of perceptual learning increases with increased training. *Vision Research*, *50*, 1928-1940.
- Jones, L. A., Kunkel, J., & Piatetski, E. (2009). Vibrotactile pattern recognition on the arm and back. *Perception*, *38*, 52-68.
- Kaczmarek, K. A., & Haase, S. J. (2003). Pattern identification and perceived stimulus quality as a function of stimulation current on a fingertip-scanned electro tactile display. *IEEE Transaction on Neural System Rehabilitation Engineering*, *11*, 9-16.
- Katz, D. (1925). *The World of Touch*. Erlbaum, Hillsdale, NJ, USA (translation published 1989, translated by L. E. Kmegeer).

- Katz, B., Dramas, F., Parseihian, G., Gutierrez, O., Kammoun, S., Brillhault, A., Brunet, L., Gallay, M., Oriola, B., Auvray, M., Truillet, P., Thorpe, S., & Jouffrais, C. (2012-a). NAVIG guidance system for the visually impaired using virtual augmented reality. *Technology and Disability, 24*, 163-178.
- Katz, B., Kammoun, S., Parseihian, G., Gutierrez, O., Brillhault, A., Auvray, M., Truillet, P., Thorpe, S., & Jouffrais, C. (2012-b). NAVIG: Augmented reality guidance system for the visually impaired. *Virtual Reality Journal, 16*, 253-269.
- Kupers, R., Fumal, A., de Noordhout, A. M., Gjedde, A., Schoenen, J., & Ptito, M. (2006). Transcranial magnetic stimulation of the visual cortex induces somatotopically organized qualia in blind subjects. *Proceedings of the National Academy of Science of the USA, 103*, 13256-13260.
- Levy-Tzedek, S., Novick, I., Arbel, R., Abboud, S., Maidenbaum, S., Vaadia, E., & Amedi, A. (2012). Cross-sensory transfer of sensory-motor information: visuomotor learning affects performance on an audiomotor task, using sensory-substitution. *Scientific Reports, 2*, 949.
- Levy-Tzedek, S., Riemer, D., & Amedi, A. (2014). Color improves “visual” acuity via sound. *Frontiers in Neuroscience, 8*, 358.
- Liu, Z., & Weinshall, D. (2000). Mechanisms of generalization in perceptual learning. *Vision Research, 40*, 97-109.
- Lopez, C. (2016). The vestibular system: balancing more than just the body. *Current Opinion in Neurology, 29*, 74-83
- Maidenbaum, S., Hanassy, S., Abboud, S., Buchs, G., Chebat, D. R., Levy-Tzedek, S., & Amedi, A. (2014). The “EyeCane”, a new electronic travel aid for the blind: Technology, behavior & swift learning. *Restorative Neurology and Neuroscience, 32*, 813-824.
- Mancini, F., Bauleo, A., Cole, J., Lui, F., Porro, C. A., Haggard, P., & Iannetti, G. D. (2014). Whole-body mapping of spatial acuity for pain and touch. *Annals of Neurology, 75*, 917-924.
- Meijer, P. B. L. (1992). An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering, 39*, 112-121.
- Menary, R. (2006). Attacking the bounds of cognition. *Philosophical Psychology, 19*, 329-324.
- Menary, R. (2007). *Cognitive integration. Mind and cognition unbounded*. Basingstoke: Palgrave.
- Meredith, M. A., & Stein, B. E. (1986). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Research, 365*, 350-354.
- Merzenich, M. M., Kaas, J. H., Sur, M., & Lin, C.-S. (1978). Double representation of the body surface within cytoarchitectonic area 3b and 1 in “SI” in the owl monkey (*Aotus trivirgatus*), *Journal of Computational Neurology, 181*, 41-74.
- Nakamura, A., Yamada, T., Goto, A., Kato, T., Ito, K., Abe, Y., Kachib, T. & Kakigic, R. (1998). Somatosensory homunculus as drawn by MEG, *NeuroImage, 74*, 377-386.
- Noë, A. (2004). *Action in Perception*. Cambridge, MA: MIT Press.
- O’Brien, J., & Auvray, M. (2016). Cognition overrides orientation dependence in tactile viewpoint selection. *Experimental Brain Research, 234*, 1885-1892.

- O'Brien, J., & Auvray, M. (2014). The process of distal attribution illuminated through studies of sensory substitution. *Multisensory Research*, 27, 421-441.
- O'Brien, J., Auvray, M., & Hayward, V. (2015). Perception of distance-to-obstacle through time-delayed tactile feedback. *Proceedings of the 2015 IEEE World Haptics Conference* (pp. 7-12).
- O'Brien, J., Hayward, V., & Auvray, M. (in preparation). Frames of reference at the tip of your finger and in the palm of your hand.
- O'Brien, J., Terekhov, A., Auvray, M., & Hayward, V. (2014). Haptic shape constancy across distance. *Haptics: Neuroscience, Devices, Modeling, and Applications Lecture Notes in Computer Science* (pp 77-84). Springer: Heidelberg, New-York, Dordrecht.
- O'Regan, J. K. (2011). *Why Red Doesn't Sound Like a Bell. Understanding the Feel of Consciousness*. Oxford: Oxford University Press.
- Parsons, L., & Shimojo, S. (1989). Perceived spatial organization of cutaneous patterns on surfaces of the human body in various positions. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 488-504.
- Parise, C. V., & Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: A study using the implicit association test. *Experimental Brain Research*, 220, 319-333.
- Pavlidou, A., Ferrè, E., & Lopez, C. (2018). Vestibular stimulation makes people more egocentric. *Cortex*, 101, 302-305,
- Pasqualotto, A., Spiller, M. J., Jansari, A. S., & Proulx, M. J. (2013). Visual experience facilitates allocentric spatial representation. *Behavioural Brain Research*, 236, 175-179.
- Petkova, V. I., Khoshnevis, M., & Ehrsson, H. H. (2011). The perspective matters! Multisensory integration in ego-centric reference frames determines full-body ownership. *Frontiers in Psychology*, 2:35.
- Pesnot-Lerousseau, J., Arnold, G., & Auvray, M. (submitted). Visualizing sounds: training-induced plasticity with a visual-to-auditory conversion device.
- Prinz, J. (2006). Putting the brakes on enactive perception. *Psyche*, 12, 1-19.
- Pritchett, L. M., & Harris, L. R. (2011). Perceived touch location is coded using a gaze signal. *Experimental Brain Research*, 213, 229-234.
- Proulx, M. J. (2010). Synthetic synaesthesia and sensory substitution. *Consciousness and Cognition*, 19,501-503.
- Proulx, M.J., Brown, D. J., Pasqualotto, A., & Meijer, P. (2014). Multisensory perceptual learning and sensory substitution. *Neuroscience and Biobehavioural Reviews*, 41, 16-25.
- Proulx, M. J., Gwinnutt, J., Dell'Erba, S., Levy-Tzedek, S., de Sousa, A. A., & Brown, D. J. (2016). Other ways of seeing: From behavior to neural mechanisms in the online "visual" control of action with sensory substitution. *Restorative Neurology and Neuroscience*, 34, 29-44.
- Proulx, M. J., Stoerig, P., Ludowig, E., & Knoll, I. (2008). Seeing 'where' through the ears: effects of learning-by-doing and long-term sensory deprivation on localization based on image-to-sound substitution. *PLoS one*, 3, e1840.
- Ptito, M., Iversen, K., Auvray, M., Deroy, O., Kupers, R. (forthcoming). The pluripotent visual cortex in congenital blindness. In F. Macpherson (Ed.). *Sensory Substitution and*



*Augmentation*, London: Proceedings of the British Academy.

- Ptito, M., Moesgaard, S. M., Gjedde, A., & Kupers, R. (2005). Cross-modal plasticity revealed by electro- tactile stimulation of the tongue in the congenitally blind. *Brain* 128, 606-614.
- Renault, A. G., Auvray, M., Parseihian, G., Miall, R. C., Cole, J., & Sarlegna, F. (2018). Does proprioception influence human spatial cognition? A study on individuals with massive deafferentation. *Frontiers in Psychology*. doi.org/10.3389/fpsyg.2018.01322
- Renier, L., Laloyaux, C., Collignon, O., Tranduy, D., Vanlierde, A., Bruyer, R., & De Volder, A. G. (2005). The Ponzo illusion using auditory substitution of vision in sighted and early blind subjects. *Perception*, 34, 857-867.
- Sekiyama, K. (1991). Importance of head axes in perception of cutaneous patterns drawn on vertical body surfaces. *Perception & Psychophysics*, 49, 481-492.
- Shore, D. I., Spry, E., & Spence, C. (2002). Confusing the mind by crossing the hands. *Cognitive Brain Research*, 14, 153-163.
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception & Psychophysics*, 73, 971-995.
- Stevens, J., & Choo, K. K. (1996). Spatial acuity of the body surface over the life span. *Somatosensory and Motor Research*, 13, 153-166.
- Stiles, N. R., & Shimojo, S. (2015). Sensory substitution. In *The Oxford Handbook of Perceptual Organization*. Oxford: Oxford University Press.
- Striem-Amit, E., & Amedi, A. (2014). Visual cortex extrastriate body-selective area activation in congenitally blind people “seeing” by using sounds. *Current Biology*, 6, 687-692.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662.
- Terekhov, A.V., & Hayward, V. (2015). The brain uses extra-somatic information to estimate limb displacement. *Proceedings of the Royal Society B*, 282, 20151661.
- Trivedi, B. (2010). Sensory hijack: rewiring brains to see with sound. *Science*, 2773, 42-45.
- van Leeuwen, C., Verstijnen, I. M., & Hekkert, P. (1999). Common unconscious dynamics underly uncommon conscious effect: A case study in the iterative nature of perception and creation In. J. S. Jordan (Ed.), *Modeling consciousness across the disciplines* (pp. 179-218). Lanham MD: University Press of America.
- Vérité, F., Bachta, W. and Morel, G. (2014). Closed Loop Kinesthetic Feedback for Postural Control Rehabilitation. *IEEE Transactions on Haptics*, 7, 150-160.
- Viaud-Delmon, I., & Warusfel, O. (2014). From ear to body: the auditory-motor loop in spatial cognition. *Frontiers in Neurosciences*, 8, 283.
- Wang, X., Zhou, Y., & Liu, Z. (2013). Transfer in motion perceptual learning depends on the difficulty of the training task. *Journal of Vision*, 13, 1-9.
- Ward, J, & Meijer, P. (2010). Visual experiences in the blind induced by an auditory sensory substitution device. *Consciousness and Cognition*, 19, 492-500.
- Ward, J., & Wright, T. (2014). Sensory substitution as an artificially acquired synaesthesia. *Neuroscience and Biobehavioral Review*, 41, 26-35.

- Wexler, M., & Held, R.M. (2005). Anticipating the three-dimensional consequences of eye movements. *Proceedings of the National Academy of Science of the USA*, *102*, 1246-1251.
- White, B. W., Saunders, F. A., Scadden, L., Bach-y-Rita, P., & Collins, C.C. (1970). Seeing with the skin. *Perception & Psychophysics*, *7*, 23-27.
- Zwiers, M. P., Van Opstal, A. J., & Paige, G. D. (2003). Plasticity in human sound localization induced by compressed spatial vision. *Nature Neuroscience*, *6*, 175-181.

## CURRICULUM VITAE - MALIKA AUVRAY

CNRS – CR1

ISIR – Institut des Systèmes Intelligents et de Robotique

Mail: [auvray@isir.upmc.fr](mailto:auvray@isir.upmc.fr)

Web: [www.malika-auvray.com](http://www.malika-auvray.com)

Tel: 0660481444

### ACADEMIC EXPERIENCE

---

**2015-Present:** CNRS Researcher (CR1), Institut des Systèmes Intelligents et de Robotique, CNRS UMR 7222.

**2008-2014:** CNRS Researcher (CR2), Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur, CNRS UPR 3251.

**2007-2008:** Post-doctoral researcher, Laboratoire Neurophysique et Physiologie, CNRS UMR 8119, supervised by Agnès Roby-Brami and Sylvain Hannequin.

**2006-2007:** Post-doctoral researcher, Centre for Philosophical Psychology, University of Antwerp, supervised by Prof. Erik Myin.

& Visiting Scholar (5 months), Department of Experimental Psychology, University of Oxford, UK, supervised by Prof. Charles Spence.

**2005-2006:** Post-doctoral researcher, Department of Experimental Psychology, University of Oxford, UK, supervised by Prof. Charles Spence.

**2004-2005:** A.T.E.R. (Associate Lecturer) in Psychology, University of Compiègne, France. Member of the Perceptual Supplementation Group, supervised by Pr Charles Lenay.

**02/2004-06/2004:** Research contract for the European Network of Excellence “Enactives Interfaces”, hired by the CNRS, UMR 8581.

**2001-2004:** Ph.D. in Psychology, in collaboration between the Laboratory of Experimental Psychology, Paris, supervised by J. Kevin O'Regan (CNRS, UMR 8581) and the company BaBeL@Stal, supervised by Olivier Mevel.

### EDUCATION

---

**2001-2004:** Ph.D. in Psychology, with honours, Ecole des Hautes Etudes en Sciences Sociales, Paris. Defended on the 20/12/2004.

Title: Immersion et perception spatiale. L'exemple des dispositifs de substitution sensorielle [Immersion and spatial perception. The example of sensory substitution devices] (297 pages). Ph.D. Thesis, EHESS, Paris.

Members of the committee: J. Kevin O'Regan (PhD supervisor), Pr Charles Lenay (President of the committee), Pr Yvette Hatwell (referee), Jérôme Dokic (referee), Pr Philippe Fuchs (committee member), Pr Michel Imbert (committee member).

**1999-2000:** M2 in Cognitive Sciences, Specialization in Psychology, with honours, Ecole des Hautes Etudes en Sciences Sociales, Paris.

Title: La cécité aux changements progressifs [Blindness to progressive changes] (95 pages). DEA, EHESS, Paris.

Supervised by: Dr J. Kevin O'Regan and Dr Pierre Jacob.

**1998-1999:** Master in Philosophy and Logic, with honours, University of Nantes.

Supervised by: Dr Nelly Robinet.

Research report: Auvray, M. (1999). Bergson et les neurosciences [Bergson and the neuroscience] (133 pages). Master, University of Nantes.

## **OTHER QUALIFICATIONS**

**2007.** Qualification to the profession of lecturer in Psychology by the French National University Council (qualification given as a function of the teaching experience and publications in the relevant discipline, necessary in order to apply for lectureships).

**2005.** Test of English for international communication (Toeic). Total: 840 (Listening: 430, reading: 410). Corresponds to the level 2+: Advanced Working Proficiency.

## **GRANTS AND FELLOWSHIPS**

---

**2018-2019:** One year post-doctoral fellowship funded by the company Oculus (PI. V. Hayward, ISIR & London University). Partners: M. Auvray (ISIR), O. Deroy, and Barry Smith (London University).

**2018:** Program Auton from the CNRS (MI, Mission for Interdisciplinarity). Renewal for the consortium Subilma (PI: F. Sarlegna, 24 K€).

**2017-2018:** Grant from the Labex SMART, project SAHMI (PI: Gilles Bailly, 186,5 K€). Partners: Y. Jansen, M. Auvray, S. Haliyo (ISIR), C. Lenay (Compiègne).

**2017:** Program Auton from the CNRS (MI). Renewal for the consortium Subilma (PI: N. Jarasse, 35 K€).

**2016:** Post-doctoral grant from the Labex SMART, in co-supervision with Nathanaël Jarasse (ISIR, UMR 7222) (beneficiary, G. Arnold, 12 months).

**2016-2019:** Grant from the French national agency for research (ANR), project: Developmental Tool Mastery (PI: A. Farne, 412, 3 K€). Partners: A. Farne, A. Roy (ImpAct, U864), F. de Vignemont (IJN, UMR 8129), M. Auvray, V. Hayward (ISIR).

**2016-2019:** ANR JCJC, I-gait (PI: W. Batcha, 235 K€). Partners: M. Auvray, N. Jarasse (ISIR), R. Reynolds (University of Birmingham).

**2016:** Program Auton from the CNRS (MI). Consortium Subilma (PI: N. Jarasse, 35 K€). Partners similar to the consortium Supplé-Sens 2014 & 2015.

**2015:** Program Auton from the CNRS (MI). Consortium Supplé-Sens (PI: M. Auvray, 35 K€). Partners similar to the consortium Supplé-Sens 2014 funded by DEFI-SENS.

**2014:** Program DEFI-SENS from the CNRS (MI). Perceptual deficiencies and multimodal supplementation. (PI: M. Auvray, 35 K€). Partners: F. Sarlegna (UMR 7287), K. O'Regan (UMR 8242), I. Fasiello (UMR 7222), P. Barone (UMR 5549), F. Matonti (UMR 7289), G.

Parseihian (UPR 7051), I. Viaud-Delmon (UMR 9912), O. Derooy (University of London), J. Cole (Poole Hospital), C. Miall (University of Birmingham).

**2013-2015:** Grant from the “Centre Francilien de l’Innovation”. Project: Elaboration and evaluation of a visuo-tactile supplementation device for blind persons (PI: M. Auvray, 23 K€). Partners: Company Caylar, Ile de France (08/2013-12/2015).

**2013:** Program DEFI-SENS from the CNRS (MI). Optimisation of sensory substitution devices using crossmodal correspondences (PI: M. Auvray, 30 K€). Partners: O. Derooy (University of London), S. Hanneton (UMR 8119), K. O’Regan (UMR 8242), I. Fasiello (UMR 7222), F. de Vignemont (UMR 8129).

**2012:** Program DEFI-SENS from the CNRS (MI). Fundamental challenges for perceptual supplementation (PI: M. Auvray, 20 K€). Partners: O. Derooy (University of London), S. Hanneton (UMR 8119), K. O’Regan (UMR 8242), V. Hayward (UMR 7222).

**2011-2014:** Grant from the French national agency for research (ANR-JCJC), project FRESCO (ANR-11-JSH2-003-1). Frames of Reference in Spatial Cognition (PI: M. Auvray, 169,94 K€, 75%). Partners: Jérôme Dokic, Frédérique de Vignemont (IJN, UMR 8129), Mehdi Ammi, Sylvain Caillou (LIMSI, UPR 3251), Charles Spence (Oxford University).

**2009-2011:** Grant from the French national agency for research (ANR-08-TECS-0011). Project NAVIG. Navigation aided by artificial vision and GNSS (PI: Jouffrais, IRIT, UMR 5505, 109 K€). Partners: S. Thorpe (CerCo, UMR 5549), M. Auvray, M. Denis, B. Katz (LIMSI, UPR 3251), Navocap, Institut des Jeunes Aveugles.

**2008-2010:** Post-doctoral Fellowship Marie Curie (Intra European Fellowship).

**2006-2007:** Post-doctoral Fellowship FWO (Belgium).

**2005-2006:** Post-doctoral Fellowship Lavoisier, French Ministry of Foreign Affairs.

**2004:** Fellowship from the Fondation de France (to develop a project based on the technologies of sensory substitution).

**2001-2004:** Ph.D. Fellowship from the A.N.R.T. (French National Association for Technical Research).

**2000-2001:** Master 2 fellowship, Ecole des Hautes Etudes en Sciences Sociales.

### **Grants obtained by PhD and post-doctoral researchers working under my supervision**

**2015-2018:** PhD grant CIFRE, in co-supervision with Jérémy Lafraire (Institut Paul Bocuse) in collaboration with the company Pernod Ricard, for Jérémy Roque (36 months).

**2015-2017:** Post-doctoral fellowship CONICET (National Scientific and Technical Research Council, Argentina), in collaboration with Claudia Arias (Cordoba, Argentina), for Laura Fernandez (24 months).

**2014-2015:** Post-doctoral fellowship from “Fondation des Aveugles de Guerre », for Gabriel Arnold (12 months).

**2013-2015:** Post-doctoral fellowship from the foundation Fyssen, for Jessica Hartcher O’Brien (24 months).

**2010-2013:** PhD grant from the university Paris 11, in collaboration with Mehdi Ammi (LIMSI, UPR 3251), for Adrien Girard (36 months).

## **ORGANISATIONS AND MANAGEMENT OF RESEARCH**

---

### **Editorial activities and organisations**

**2016.** Organization of the Symposium *Sensory Substitution* at *International Congress of Psychology*, 24-29 July, Yokohama, Japan.

**2014.** Editor of the special issue *Sensory Substitution* in the journal *Multisensory Research*.

**2014.** Co-editor of *Haptics: Neuroscience, Devices, Modeling, and Applications*, Springer.

Co-presidency of the program for the conference *Eurohaptics 2014*, 24-26 June, Versailles, France. The conference gave rise to the abovementioned journal special issue *Haptics*.

**2014.** Member of the Copil of the experimental platform at the DEC.

**2002.** Organization of the *Third Workshop on Multimodal Interaction in Perception*, 28-29 October, Paris, France.

### **Selection committee**

**2016.** Selection committee for the lecturer positions « 16 MCF0065, cognition et action » and « 16 MCF0125, mémoire et conscience » University Paris 5.

**2015.** Selection committee for the lecturer positions « 16 MCF0120, perception et action » and « 16 MCF1826, mémoire humaine » University Paris 5.

## **PARTICIPATION TO THE PRESS, TV, AND RADIO FOR THE GENERAL PUBLIC**

### **Participation to journals for the general public**

Auvray, M., & Dauphin, B. (2015). Les sens cachés du vin. *Revue des Œnologues*, 157, p. 21.

Dauphin, B., & Auvray, M. (2014). Le vin, une expérience multisensorielle. *Revue des Œnologues*, 41, 17-18.

Auvray, M., & O'Regan, J. K. (2003). Voir avec les oreilles: Enjeux de la substitution sensorielle. *Pour la Science*, 39, 30-35.

### **TV, radio, and public events**

Participation to the documentary « Les forces du toucher » in the series « Nos cinq sens », broadcasted on Canal+, Planete+, TV5Monde, TV5Québec & RTBF in January 2018. Realisation: Philippe Allante. Filming the 8 & 28 March 2017.

Participation to the audio documentary « Il faut croire pour voir » broadcasted on France Culture the 8 Mars 2017, Realisation: Tony Hayere, recorded the 16 January 2017.

Auvray, M. (2017). Les dispositifs de suppléance perceptive à destination des personnes non voyantes. *Conférence – Débat « Sciences et Cécités »*, *Voir avec les oreilles, voir avec les doigts*, 8 April, Blois, France.

Participation to the show 'On n'est pas que des cobayes' on sensory substitution, filming in November 2013, broadcasted the 31 January 2014.

Bloomsbury Festival, 19<sup>th</sup> October 2013, Centre for the Study of the Censes, Institute of Philosophy, London.

Presentation of posters during "Forum des Laboratoires en Sciences Cognitives d'Ile de France" (2002 and 2003)

Participation to the days “Fête de la science” at Compiègne (2003) and Paris (2002).

## **STUDENTS, PHD, AND POST-DOC SUPERVISION**

---

- 2018-2020. Shared supervision (50%) with Vincent Hayward (ISIR) of a post-doctoral researcher (Xavier Job), ANR project: « Developmental tool mastery ».
- 2018-2020. Shared supervision (50%) with Vincent Hayward (ISIR) of a post-doctoral researcher (Louise Kirsch). Title of the project: «The processes of exteriorization and embodiment in tool use ».
- 2018-2019. Shared supervision (1/3) with Vincent Hayward (ISIR) and Ophelia Derooy (London) of a post-doctoral researcher (Antonio Cataldo), funded by Oculus VR. Title of the project: « Distal projection of tactile sensations from oculomotor signals »
- 2017-2018. Supervision of a student in Master 2 in Psychology at the university Paris 5 (Yann Kechabia). Title of the research: « Retour de proprioception artificiel par stimulation vibrotactile ».
- 2017-2018. Supervision of a student in Master 1 of the Cogmaster (Nathan Risch). Title of the research: « Existe-t-il une différence de toucher entre soi et autrui ? »
- 2017-2018. Supervision of a student in Master 1 of the Cogmaster (Daphne Parmentier). Title of the research: « Différences individuelles en substitution sensorielle visuo-auditive ».
- 2017-2018. Supervision of 6 students in third year of Polytechnic School. Title of the research: « Substitution sensorielle: voir par le toucher ».
- 2015-2018: Shared supervision (50%) with Jérémy Lafraire (Institut Paul Bocuse) of a PhD student (Jérémy Roque) CIFRE project in collaboration with Pernod Ricard. Title of the PhD: « Perception multimodale des boissons alcoolisées: le cas particulier de la perception de fraîcheur ».
- 2016-2017. Supervision of a student in M1 of the Cogmaster (Sacha Yesislatay). Title of the research : « Cognition spatiale chez les patients atteints d'autisme ».
- 2015-2017: Shared supervision (50%) with Claudia Arias (Cordoba, Argentina) of a post-doctoral researcher (Laura Fernandez) with a post-doctoral Grant CONICET (National Scientific and Technical Research Council).
- 2016-2017. Supervision of a student in M2 in Psychology at the University Paris 5 (Chenyang Wang). Title of the research: « Conscience corporelle et cognition spatiale ».
- 2016-2017. Supervision of a student in M2 in Psychology at the University Paris 5 (Irène Rabotin). Title of the research: « Le rôle de l'expérience visuelle pour l'appropriation d'un dispositif de substitution sensorielle ».
- 2016-2017. Supervision of a student in M1 in Psychology at the University Paris 8 (Sandra Inard). Title of the research: « Etude des facteurs psychologiques déterminant les cadres de références utilisés ».
- 2015-2016. Supervision of a student in M2 of the Cogmaster (Jacques Pesnot-Lerousseau). Title of the research: « Etude des processus visuels et auditifs impliqués dans l'utilisation du dispositif de substitution sensorielle The vOICe ».
- 2012-2016: Supervision of a post-doctoral researcher (Gabriel Arnold) with grants from the ANR FRESCO (2 years), post-doctoral fellowship from “Fondation des Aveugles de

- Guerre” (1 year), fellowship from the Labex SMART (1 year in co-supervision with Nathanael Jarasse, 50%).
- 2014-2015. Shared supervision with Fabrice Sarlegna (CNRS UMR 7287), of a student (Alix Renault) in M2 STAPS. Title of the research: « Influence de la proprioception sur la représentation de l'espace ».
- 2013-2015. Supervision of a post-doctoral researcher (Jessica Hartcher O'Brien), with a grant from Fyssen.
- 2013-2014: Shared supervision with Frédérique de Vignemont (CNRS UMR 8129) of a student in M2 of the PhilMaster (Kahina Ouibrahim). Title of the research: « Le sens du self ».
- 2013-2014: Supervision of a student in M2, CNA (Systèmes complexes: Cognitions naturelle et artificielle) (Fatma Elleuch). Title of the research: « Inversion ou renforcement de la correspondance intermodale entre la fréquence auditive et l'élévation visuelle ».
- 2013-2014. Supervision of three students (Erika Kaiser, Marie Lorenz, Frédéric Albert) in Master 2 Ingénierie de l'innovation, Valorisation de la recherche et transfert de compétences. Title of the research: « Matrice de suppléance perceptive visuo-tactile pour personnes non-voyantes ».
- 2012-2013: Supervision of a student in M2, Cogmaster (Kahina Ouibrahim). Title of the research: « l'effet des facteurs de dominance sur la prise de perspective spatiale ».
- 2011-2012: Shared supervision with Sylvain Hanneton (CNRS, UMR 8119) of a student in M2, VHMA, Paris 5 (Philippe Herquel). Title of the research: « L'apprentissage de la pratique d'un jeu vidéo purement sonore ».
- 2010-2014: Shared supervision (50%) with Mehdi Ammi (CNRS, UPR 3251) of a PhD student from Paris 11 (Adrien Girard). Title of the PhD: « Interactions visio-haptiques pour la coordination et la conscience situationnelle partagée en environnement virtuel collaboratif ».
- 2009-2011: Shared supervision with Michel Denis (CNRS, UPR 3251) of a post-doctoral researcher (Mathieu Gallay), grant from the ANR Navig.
- 2010-2011: Shared supervision with Sylvain Hanneton (CNRS, UMR 8119) of a student in M2 VHMA, Paris 5 (Vincent Forma). Title of the research: « Les oreilles sur la main: l'étude de l'apprentissage d'un nouveau couplage audio-moteur lors d'une tâche de pointage en environnement sonore 3D ».
- 2009-2010: Supervision of a student in M2 Ergonomie et Facteurs Humains, Paris 11 (Lucie Brunet). Title of the research: « Etude des stratégies et des besoins des personnes non-voyantes lors de la navigation pour la conception d'un dispositif d'aide performant et accepté ».
- 2009-2010: Supervision of a student in M2 ATIAM, Paris 6 (Lourdes Garcia). Title of the research: « Étude des représentations spatiales en environnement sonore ».
- 2008-2009: Shared supervision with Mark Wexler (CNRS, UMR 8158) of a student in M2 of the Cogmaster. Title of the research: « Le rôle de l'activité motrice dans le raisonnement spatial ».
- 02/2005-06/2005: Shared supervision with Charles Lenay (Université de Technologie de Compiègne) of two students in 4<sup>th</sup> year of engineering school. Title of the research: « Les interactions perceptives en environnement minimaliste ».



02/2004-06/2004: Shared supervision with Sylvain Hanneton (CNRS, UMR 8119) of a student in M1 STAPS, Paris 5.

10/2003-01/2004: Shared supervision with Charles Lenay (Université de Technologie de Compiègne) of four students in 3<sup>ème</sup> year of engineering school.

04/2003-07/2003: Shared supervision, with J. Kevin O'Regan of a student in 3<sup>ème</sup> year of Polytechnic school.

06/2003: Supervision of three students in 1st year of Psychology, Université Paris 5.

### Participation to PhD committees

- Lise Obeika, “Interplay between multisensory integration and social interaction in auditory space”. Defended the 29 November 2017. Isabelle Viaud-Delmon (PhD supervisor), Ana Tajadura-Jiménez (referee), Stefan Glasauer (referee), Guenther Knoblich (examiner), Malika Auvray (examiner), Giandomenico Iannetti (examiner).

- Julien Laroche, « Le processus d’interaction comme attracteur dans la coordination temporelle entre les personnes ». Defended the 24 June 2013. Pr Eric Brangier (directeur de thèse), Anna-Maria Berardi (PhD supervisor), Pr Olivier Gapenne (referee), Pr Maya Gratier (referee), Malika Auvray (examiner).

- Adrien Girard, « Interactions visio-haptiques pour la coordination et la conscience situationnelle partagée en environnement virtuel collaboratif ». Defended the 12 June 2014. M. Ammi (PhD supervisor), M. Auvray (PhD co-supervisor), J.-M. Burkhardt (referee), G. Moreau (referee), A. Lécuyer (examiner), A. Vilnat (examiner), G. Zachmann (e examiner), Y. Bellik (examiner).

### V. Teaching

Total: 480 hours. Including 54h CM, 380h TD, 85h tutoring

Institution	Title	Public	Years
Département d'Etudes cognitives (DEC – ENS)	Cognition incarnée	Master (année 2)	2017-2018 (12h) 2016-2017 (12 h) 2015-2016 (12 h) 2014-2015 (12 h) 2013-2014 (12 h) 2012-2013 (12 h)
UPMC	Perception multisensorielle appliquées aux technologies de réhabilitation	Master (année 2) Filière MSR	2017-2018 (8h)
Université Paris 5	Substitution sensorielle	Master (année 2) Filière PCFA	2017-2018 (6h) 2016-2017 (4 h) 2015-2016 (2 h)
	Psychologie Expérimentale	Licence (année 1) de psychologie	2003-2004 (48 h) 2002-2003 (48 h)
	Méthodologies en psychologie (tutorat)	Licence (année 1) de psychologie	2002-2003 (45 h)

<b>ENSCI – les ateliers</b>	Perception multisensorielle et synesthésie	1 <sup>ère</sup> à 5 <sup>ème</sup> année de l'école de design	2017-2018 (10h)
<b>Université Paris 11</b>	Approfondissements en Psychologie Cognitive	Master (année 2)	2009-2010 (9 h) 2008-2009 (9 h)
<b>Université de Technologie de Compiègne</b>	Sciences cognitives et naturalisation de l'esprit	1 <sup>ère</sup> à 5 <sup>ème</sup> année de l'école d'ingénieur	2004-2005 (68 h) ½ ATER Psychologie
	Apprentissage et perception	1 <sup>ère</sup> à 5 <sup>ème</sup> année de l'école d'ingénieur	2002-2003 (8 h) 2001-2002 (8 h)
	Psychoacoustique	5 <sup>ème</sup> année de l'école d'ingénieur	2002-2003 (4 h) 2001-2002 (4 h)
<b>Université Paris 13</b>	Perception et cognition	Licence (année 1) de psychologie	2002-2003 (48 h) 2001-2002 (35 h)
<b>ESIGELEC de Rouen</b>	Perception et action	5 <sup>ème</sup> année de l'ESIGELEC	2002-2003 (2 h)
<b>Ecole Normale Supérieure</b>	La substitution sensorielle	1 <sup>ère</sup> année de l'ENS	2002-2003 (2 h)
<b>Université de Nantes</b>	Méthodologie en philosophie (tutorat)	Licence (année 1) de philosophie	1998-1999 (20 h) 1997-1998 (20 h)

## PUBLICATIONS

---

h-index: 19 / Citations: 1921 (Google Scholar)

The names of the PhD and master students I have supervised are underlined and in bold. The names of the post-doctoral researchers I have supervised are underlined.

### ARTICLES

1. Roque, J., Lafraire, J., Spence, C., & Auvray, M. (in press). The influence of audiovisual perceptual features on the categorization of freshness in beverages. *Journal of Sensory Studies*.
2. **Renault**, A. G., Auvray, M., Parseihian, G., Miall, R. C., Cole, J., & Sarlegna, F. (2018). Does proprioception influence human spatial cognition? A study on individuals with massive deafferentation. *Frontiers in Psychology*. doi.org/10.3389/fpsyg.2018.01322
3. Arnold, G., & Auvray, M. (2018). Tactile recognition of visual stimuli: Specificity versus generalization of perceptual learning. *Vision Research*. doi.org/10.1016/j.visres.2017.11.007
4. **Brunet**, L. Darses, F., & Auvray, M. (2018). Strategies and needs of blind pedestrians during urban navigation, *Le Travail Humain*. 81, 141-171.
5. **Roque**, J., Auvray, M., & Lafraire, J. (2018). Investigating the cognitive mechanisms underlying freshness perception: a research review. *Frontiers in Psychology*, 8, 2360. doi.org/10.3389/fpsyg.2017.02360
6. Arnold, G., Spence, C., & Auvray, M. (2017). A unity of the self or a multiplicity of locations? How the graphesthesia task sheds light on the role of spatial perspectives in

- bodily self-consciousness. *Consciousness & Cognition*, 56, 100-114. doi: 10.1016/j.concog.2017.06.012
7. Arnold, G., Pesnot-Lerousseau, J. & Auvray, M. (2017). Individual differences in sensory substitution. *Multisensory Research*, 6, 579-600. doi: 10.1163/22134808-00002561
  8. Arnold, G. & Auvray, M. (2017). The graphesthesia paradigm: drawing letters on the body to investigate the embodied nature of spatial perspective taking. *I-Perception*, 8, 1-5. doi: 10.1177/2041669517690163
  9. Deroy, O., Fasiello, I., Hayward, V., & Auvray, M. (2016). Differentiated audio-tactile correspondences in sighted and blind individuals. *Journal of Experimental Psychology: Human Perception and Performance*, 42, 1204-1214. doi:10.1037/xhp000015
  10. O'Brien, J., & Auvray, M. (2016). Cognition overrides orientation dependence in tactile viewpoint selection. *Experimental Brain Research*, 234, 1885-1892. doi 10.1007/s00221-016-4596-6
  11. Arnold, G., Spence, C., & Auvray, M. (2016). Taking someone else's spatial perspective: Natural stance or effortful decentring? *Cognition*, 148, 27-33. doi:10.1016/j.cognition.2015.12.006
  12. Hanneton, S., Herquel, P., & Auvray, M. (2015). Intermodal recoding of a video game: Learning to process signals for motion perception in a pure auditory environment. *International Journal of Adaptive Control and Signal Processing*, 29, 1475-1483. doi: 10.1002/acs.2549
  13. O'Brien, J., & Auvray, M. (2014). The process of distal attribution illuminated through studies of sensory substitution. *Multisensory Research*, 27, 421-441. doi: 10.1163/22134808-00002456
  14. Auvray, M., & Harris, L.R. (2014). The state of the art of sensory substitution. *Multisensory Research*, 27, 265-269. doi: 10.1163/22134808-00002464
  15. Arnold, G., & Auvray, M. (2014). Perceptual learning: Tactile letter recognition transfers across body surfaces. *Multisensory Research*, 27, 71-90. doi: 10.1163/22134808-00002443
  16. Deroy, O., & Auvray, M. (2012). Reading the world through sensory substitution devices. *Frontiers in Theoretical and Philosophical Psychology*, 3, 457. doi: 10.3389/fpsyg.2012.00457
  17. Auvray, M., & Rohde, M. (2012). Perceptual crossing: The simplest online paradigm. *Frontiers in Human Neuroscience*, 6, 181. doi: 10.3389/fnhum.2012.00181
  18. Simard, J., Ammi, M., & Auvray, M. (2012). Collaborative strategies for the search of 3D targets in molecular environments. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 6, 1555-1565. doi: 10.1155/2012/546521
  19. Auvray, M., Hoellinger, T., Hanneton, S., & Roby-Brami, A. (2011). Perceptual weight judgments when viewing one's own and other's movements under minimalist conditions of visual presentation. *Perception*, 40, 1081-1103. doi:10.1068/p6879
  20. Auvray, M., Gallace, A., & Spence, C. (2011). Tactile short term memory for stimuli presented on the fingertips and across the rest of the body surface. *Attention, Perception, & Psychophysics*, 73, 1227-1241. doi: 10.3758/s13414-011-0098-6

21. Hanneton, S., Auvray, M., & Durette, B. (2010). The Vibe: A versatile vision-to-audition sensory substitution device. *Applied Bionics and Biomechanics*, 7, 269-276. doi: 10.1080/11762322.2010.512734
22. Auvray, M., Myin, E., & Spence, C. (2010). The sensory-discriminative and affective-motivational aspects of pain. In F. McGlone (Ed.) *Touch, Temperature, Pleasure, and Pain*, special issue of *Neuroscience and Biobehavioral Reviews*, 34, 214-223. doi:10.1016/j.neubiorev.2008.07.008
23. Auvray, M., & Myin, E. (2009). Perception with compensatory devices. From sensory substitution to sensorimotor extension. *Cognitive Science*, 33, 1036-1058. doi: 10.1111/j.1551-6709.2009.01040.x
24. Auvray, M., Lenay, C., & Stewart, J. (2009). Perceptual interactions in a minimalist environment. *New Ideas in Psychology*, 27, 79-97. doi: 10.1016/j.newideapsych.2007.12.002
25. Auvray, M., Gallace, A., Hartcher-O'Brien, J., Tan, H. Z., & Spence, C. (2008). Tactile and visual distractors induce change blindness for tactile stimuli presented on the fingertips. *Brain Research*, 1213, 111-119. doi:10.1016/j.brainres.2008.03.015
26. Auvray, M., & Spence, C. (2008). The multisensory perception of flavour. *Consciousness & Cognition*, 17, 1016-1031. doi:10.1016/j.concog.2007.06.005
27. Auvray, M., Gallace, A., Tan, H. Z., & Spence, C. (2007). Crossmodal change blindness between vision and touch. *Acta Psychologica*, 126, 79-97. doi:10.1016/j.actpsy.2006.10.005
28. Auvray, M., Philipona, D., O'Regan, J. K., & Spence, C. (2007). The perception of space and form recognition in a simulated environment: The case of minimalist sensory-substitution devices. *Perception*, 36, 1736-1751. doi:10.1068/p5800
29. Auvray, M., Hanneton, S., & O'Regan, J. K. (2007). Learning to perceive with a visuo-auditory substitution system: Localization and object recognition with The Voice. *Perception*, 36, 416-430. doi:10.1068/p5631
30. Gallace, A., Auvray, M., & Spence, C. (2007). The modulation of haptic line bisection using a visual illusion and optokinetic stimulation. *Perception*, 36, 1003-1018. doi:10.1068/p5457
31. Gallace, A., Auvray, M., Tan, H. Z., & Spence, C. (2006). When visual transients impair tactile change detection: A novel case of crossmodal change blindness? *Neuroscience Letters*, 398, 280-285. doi:10.1016/j.neulet.2006.01.009
32. Auvray, M., Hanneton, S., Lenay, C., & O'Regan, J. K. (2005). There is something out there: Distal attribution in sensory substitution, twenty years later. *Journal of Integrative Neuroscience*, 4, 505-521. doi: 10.1142/S0219635205001002
33. Auvray, M. (2003). Bergson, une théorie sensorimotrice de la perception. *Psychologie et Histoire*, 4, 61-100.
34. Auvray, M., & O'Regan, J. K. (2003). L'influence des facteurs sémantiques sur la cécité aux changements progressifs dans les scènes visuelles. *L'Année Psychologique*, 103, 9-32. doi: 10.3406/psy.2003.29621

## BOOK CHAPTERS

1. Ptito, M., Iversen, K., Auvray, M., Deroy, O., Kupers, R. (forthcoming). The pluripotent visual cortex in congenital blindness. In F. Macpherson (Ed.). *Sensory Substitution and Augmentation*, London: Proceedings of the British Academy.
2. Auvray, M., & Farina, M. (2017). Patrolling the boundaries of synaesthesia: A critical appraisal of transient and artificially-acquired forms of synaesthetic experiences. In O. Deroy (Ed.). *Synaesthesia: Philosophical & Psychological Challenges* (pp. 248-274), Oxford: Oxford University Press. doi: 10.1093/oso/9780199688289.003.0013
3. O'Brien, J., Auvray, M., & Hayward, V. (2015). Perception of distance-to-obstacle through time-delayed tactile feedback. *Proceedings of the 2015 IEEE World Haptics Conference* (pp. 7-12). doi: 10.1109/WHC.2015.7177683
4. Auvray, M., & Deroy, O. (2015). How do synesthetes experience the world? In M. Matthen (Ed.), *Oxford Handbook of Philosophy of Perception* (pp. 640-658), Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199600472.013.027
5. O'Brien, J., Terekhov, A., Auvray, M., & Hayward, V. (2014). Haptic shape constancy across distance. *Haptics: Neuroscience, Devices, Modeling, and Applications Lecture Notes in Computer Science* (pp 77-84). Springer: Heidelberg, New-York, Dordrecht. Proceedings of Eurohaptics, 24-26 June, Versailles, France. doi: 10.1007/978-3-662-44193-0\_11.
6. Deroy, O., & Auvray, M. (2014). A crossmodal perspective on sensory substitution. In S. Biggs, M. Matthen, & D. Stokes (Eds.), *Perception and Its Modalities* (pp. 327-349), Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199832798.003.0014
7. Spence, C., Auvray, M., & Smith, B. (2014). Confusing tastes with flavours. In S. Biggs, M. Matthen, & D. Stokes (Eds.), *Perception and Its Modalities* (pp. 247-276), Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199832798.003.0011
8. Gallay, M., Denis, M., & Auvray, M. (2013). Navigation assistance for blind pedestrians: Guidelines for the design of devices and implications for spatial cognition. In T. Tenbrink, J. Wiener, & C. Claramunt (Eds.), *Representing space in cognition: Interrelations of behaviour, language, and formal models*, Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199679911.003.0011
9. Tarroux, P., & Auvray, M. (2012). La perception visuelle. In M. Denis (Ed.), *La Psychologie Cognitive* (pp. 39-69), Paris: Editions de la Maison des Sciences de l'Homme.
10. Auvray, M., & Fuchs, P. (2007). Perception, immersion et interactions sensorimotrices en environnement virtuel. In A. Grumbach & E. Klinger (Eds.), *Réalité Virtuelle et Cognition* (pp. 23-36), numéro spécial de *Intellectica*.
11. Auvray, M. (2006). Remplacer un sens par un autre: La suppléance perceptive. In P. Fuchs (Ed.), *Le Traité de la Réalité Virtuelle III*, Vol.1, L'Homme et l'Environnement Virtuel (pp. 173-188). Paris: Les Presses de l'Ecole des Mines.
12. Auvray, M. (2006). Théories de la perception: Le monde comme mémoire externe. In P. Fuchs (Ed.), *Le Traité de la Réalité Virtuelle III*, Vol.1, L'Homme et l'Environnement Virtuel (pp. 209-228). Paris: Les Presses de l'Ecole des Mines.

## OTHER PUBLICATIONS

1. Katz, B., Kammoun, S., Parseihian, G., Gutierrez, O., Brilhault, A., Auvray, M., Truillet, P., Thorpe, S., & Jouffrais, C. (2012). NAVIG: Augmented reality guidance system for the visually impaired. *Virtual Reality Journal*, 16, 253-269. doi: 10.1007/s10055-012-0213-6.
2. Katz, B., Dramas, F., Parseihian, G., Gutierrez, O., Kammoun, S., Brilhault, A., **Brunet**, L., Gallay, M., Oriola, B., Auvray, M., Truillet, P., Thorpe, S., & Jouffrais, C. (2012). NAVIG guidance system for the visually impaired using virtual augmented reality. *Technology and Disability*, 24, 163-178. doi: 10.3233/TAD-2012-0344
3. Kammoun, S., Parseihian, G., Gutierrez, O., Brilhault, A., Serpa, A., Raynal, M., Oriola, B., Macé, M., Auvray, M., Denis, M., Thorpe, S., Truillet, P., Katz B., Jouffrais, C., (2012). Navigation and space perception assistance for the visually impaired: The NAVIG project. *IRBM, Ingénierie et Recherche Biomédicale*, 33, 182-189. doi: 10.1016/j.irbm.2012.01.009
4. Gauchou, H., & Auvray, M. (2007). L'inhibition comme absence de représentation: L'exemple de la cécité aux changements. *Le Journal des Psychologues*, 244, 29-34. doi: 10.3917/jdp.244.0029
5. Auvray, M., Lenay, C., O'Regan, J. K., & Lefèvre, J. (2005). Suppléance perceptive, immersion et informations proprioceptives. *Arobase*, 1, 94-114.
6. Khatchatourov, A., & Auvray, M. (2005). L'outil modifie-t'il la perception ou la rend-il possible ? *Arobase*, 1, 65-70.
7. Auvray, M. (2005). La suppléance perceptive: Implications sur les théories de la perception. *VOIR*, n° 30-31. Les aspects culturels de la vision et les autres modalités perceptives III. L'ouïe, 2-13.

## MANUSCRIPTS SUBMITTED AND IN PREPARATION

Arnold, G., Sarlegna, F., Fernandez, L., & Auvray, M. (submitted). The physical body plays a different role in locating and orienting the experienced self in deafferented patients and in controls. *Frontiers in Neurorobotics*.

**Pesnot-Lerousseau**, J., Arnold, G., & Auvray, M. (submitted). Visualizing sounds: training-induced plasticity with a visual-to-auditory conversion device.

O'Brien, J., Hayward, V., & Auvray, M. (in preparation). Frames of reference at the tip of your finger and in the palm of your hand.

Arnold, G. & Auvray, M. (in preparation). Blind and sighted people differ in the way they mentally represent space from different perspectives.

**Roque**, J., Lafraire, J., & Auvray, M., (in preparation). Audiovisual crossmodal correspondences: the case of carbonation in beverages.

**Risch**, N., Arnold, C., & Auvray, M. (in preparation). The numbness illusion influences how you touch yourself.

## COMMUNICATIONS

*Communications in international conferences with proceedings are indicated with an asterisk.*

- \* Auvray, M., & Arnold, G., Pesnot-Lerousseau, J. (2018). Training-induced plasticity with a visual-to-auditory conversion system. Seeing the thunder while still hearing it. *International Multisensory Research Forum*. 14-17 June, Toronto, Canada.
- \* Roque, J., Lafraire, J., & Auvray, M. (2018). Audiovisual crossmodal correspondences between bubbles' size and pouring sounds' pitch in carbonated beverages. *International Multisensory Research Forum*. 14-17 June, Toronto, Canada (Poster).
- Auvray, M. (2018). *Taxonomizing the sense modalities without circularity*. *Workshop: The Philosophy of Multisensory Perception*, 14 June, Toronto, Canada.
- \* Auvray, M., & Arnold, G., & Pesnot-Lerousseau, J. (2018). Seeing the thunder while still hearing it. Functional plasticity with sensory substitution devices. *European Workshop on Imagery and Cognition*. 7-9 June, Padova, Italy.
- \* Auvray, M., & Arnold, G. (2017). Differences in the way sighted, early, and late-blind people mentally represent tactile stimuli from inside or from outside the body. *4<sup>th</sup> International Symposium Low Vision and the Brain*, 24-26 November, Berlin, Germany.
- Auvray, M. (2017). Les processus spatiaux impliqués dans la substitution sensorielle, *Séminaire du Neuroscience Research Centre of Lyon, INSERM U1028, CNRS UMR529*, 19 June, Lyon (Bron).
- Auvray, M. (2017). Multisensory and spatial processes in sensory substitution. *Séminaire du Laboratoire de Neurosciences Intégratives et Adaptatives*, UMR 7260, 24 October, Marseille
- Auvray, M. (2017). Multisensory and spatial processes in sensory substitution, *Séminaire du Laboratoire de Psychologie*, 19 June, Boulogne-Billancourt, France.
- \* Roque, J., Lafraire, J., & Auvray, M. (2017). Investigating the multisensory perception of freshness in beverages: the case of audio-visual interactions. *International Multisensory Research Forum*, 19-22 May, Nashville, USA.
- \* Auvray, M., & Arnold, G. (2016). Taking into account individual differences in multisensory processing between sighted and visually impaired people to improve sensory substitution. *Israeli Society for Neuroscience*, 4-6 December, Eilat, Israel.
- \* Arnold, G., & Auvray, M. (2016). The influence of blindness and short-term visual deprivation on spatial cognition. *Israeli Society for Neuroscience*, 4-6 December, Eilat, Israel (poster).
- Auvray, M. (2016). Multisensory perception and spatial cognition: Taking into account individual differences to improve sensory substitution. *Workshop Computational Touch*, 8-9 September, Paris, France.
- \* Arnold, G., & Auvray, M. (2016). Design and use of sensory substitution devices: differences between sighted and visually impaired people. *International Congress of Psychology*, 24-29 July, Yokohama, Japan.
- \* Auvray, M., & Arnold, G. (2016). Perceiving tactile objects from inside or from outside the body: sighted and visually impaired people differ in the way they mentally represent space

from different perspectives. *Association of the Scientific Study of Consciousness*, 13-18 June, Buenos Aires, Argentina (poster).

- \* Auvray, M., & Arnold, G. (2016). Learning to use sensory substitution devices: individual specificities and differences between sighted and visually impaired people. *European Workshop on Imagery and Cognition*, 6-8 June, Paris, France.
  - \* Arnold, G., & Auvray, M. (2016). Blind and sighted people differ in the way they mentally represent space from different perspectives. *European Workshop on Imagery and Cognition*, 6-8 June, Paris, France (poster).
  - \* Roque, J., Auvray, M., Garrel, C., & Lafraire, J. (2016). The multisensory perceptions of flavour and freshness: a new perspective. *British Feeding and Drinking Group*, 08 April, London, UK.
- Auvray, M. (2016). Replacing one modality by another: Learning to use sensory substitution devices. *Séminaire de l'Institut des Sciences du Mouvement*, 03 March, Marseille, France.
- Auvray, M., & Deroy, O. (2015). Audio-tactile crossmodal correspondences in sighted, early, and late blind people. *Seminar of the Department of Philosophy*, 18 December, PUC: Rio de Janeiro, Brasil.
- Auvray, M. (2015). Where I am? Studies on spatial perspective in the tactile sensory modality. *Seminar of the Department of Philosophy*, 14 December, PUC: Rio de Janeiro, Brasil.
- Auvray, M. (2015). Multisensory perception. *Séminaire de Recherche de l'Institut Paul Bocuse*, 13 November, Lyon, France.
- \* Hartcher-O'Brien, J., Auvray, M., & Hayward, V. (2015). Perception of distance-to-obstacle through time-delayed tactile feedback. *World Haptics Conference*, 22-26 June, Evanston, IL, USA.
  - \* Arnold, G., Hartcher-O'Brien, J., Hayward, V., Spence, C., & Auvray, M. (2015). Where am I? Investigating which perspectives are taken on ambiguous tactile symbols. *16<sup>th</sup> International Multisensory Research Forum*, 13-16 June, Pisa, Italy.
  - \* Hartcher-O'Brien, J., Auvray, M., & Hayward, V. (2015). Access to visual distance via a time-delayed tactile feedback device. *16<sup>th</sup> Multisensory Research Forum (IMRF)*, 13-16 June, Pisa, Italy.
  - \* Beaudry-Richard, A., Harrar, V., Auvray, M., Spence, C., Kupers, R., & Ptito, M. (2015). The multisensory substitution device: replacing vision with multisensory perception *16<sup>th</sup> Multisensory Research Forum (IMRF)*, 13-16 June, Pisa, Italy (poster).
- Auvray, M. (2015). Sensory substitution and augmentation in visual rehabilitation. *Seminar of the Department of Philosophy*, 20 April, PUC: Rio de Janeiro, Brasil.
- Auvray, M. (2015). Changing spatial and mental perspective: Studies in touch. *Seminar of the Department of Philosophy*, 17 April, PUC: Rio de Janeiro, Brasil.
- Arnold, G., & Auvray, M. (2014). Perceptual learning, crossmodal interactions, and spatial reference frames: Implications for sensory substitution. *Journée Robotique et Neuroscience*, 17 December, Paris, France.
- Sarlegna, F., & Auvray, M. (2014). Déficience perceptive et suppléance multimodale. *Technologies appliquées au handicap: de la compensation à la participation, DVS Defisens*, 26 September, Paris, France.



- \* Auvray, M., Deroy, O., Fasiello, I., & Hayward, V. (2014). Crossmodal correspondences between pitch and tactile movement in sighted, early, and late blind people. *European Society for Philosophy and Psychology*, 16-19 September, Noto, Sicily.
- Arnold, G., & Auvray, M. (2014). The left-side bias in tactile recognition is modulated by the adopted reference frame. *4<sup>th</sup> North Sea Meeting on Brain Asymmetries*, 10-14 September, Durham, UK (poster).
- \* **Girard**, A., Auvray, M., Ammi, M. (2014). Collaborative metaphor for haptic designation in complex 3D environments. *ACM Symposium on Applied Perception* (proceedings, pp. 31-37, doi: 10.1145/2628257.2628262), 08-09 August, Vancouver, Canada.
- \* Hartcher-O'Brien, J., Terekhov, A., Auvray, M., & Hayward, V. (2014). Haptic shape constancy across distance. *Eurohaptics*, 24-26 June, Versailles, France (poster).
- \* Arnold, G., & Auvray, M. (2014). Tactile symbol recognition: Specificity versus generalization of learning. *15<sup>th</sup> Multisensory Research Forum (IMRF)*, 11-14 June, Amsterdam, Netherlands (poster).
- \* Hartcher -O'Brien, J., & Auvray, M. (2014). The bi-stable palm: clockwise versus counter-clockwise tactile motion perception. *15<sup>th</sup> Multisensory Research Forum (IMRF)*, 11-14 June, Amsterdam, Netherlands (poster).
- Auvray, M. (2014). Questioning the use of sensory substitution devices at the behavioral, neurophysiological and phenomenal levels. *Seminar of the Paris Consciousness/Self-consciousness [PaCS]*, 31 March, Institut Jean Nicod, Paris, France.
- Auvray, M. (2014). Learning to use sensory substitution devices. *Seminar Paris Ouest*, 19 March, Nanterre, France.
- \* Auvray, M. (2014). Replacing one sense by another: Learning to use sensory substitution devices. *Symposium on Cerebral visual injury: dysfunction, plasticity, restoration and rehabilitation*; held during the *4th International Scientific Conference Restauración Neurológica*, 5-7 March, La Habana, Cuba.
- \* Auvray, M. (2014). Replacing one sense by another: learning to use a sensory substitution device. *Neuroeng2014*, 2-3 February, Adelaide, Australia.
- \* Auvray, M., & Arnold, G. (2014). Perceptual learning of tactile letters transfers across body surfaces. *Australasian Neuroscience Society (ANS)*, 28-31 January, Adelaide, Australia (poster).
- Auvray, M. (2014). L'apprentissage des dispositifs de suppléance perceptive. *Colloque Perception Augmentée et Multimodale*, 15 January, Paris, France.
- Arnold, G., & Auvray, M. (2013). Adopting and changing reference frames when interpreting tactile stimulation received on the body surface. *Symposium: vision, action and concepts: Behavioural and neural basis of embodied perception and cognition*, 28-30 October, Lille (poster).
- Auvray, M. (2013). Learning to use sensory substitution devices. *2nd HCC Summer School*, 17-19 July, Ircam, Paris.
- Auvray, M. (2013). Spatial cognition and sensory substitution. *Research Conference on Sensory Substitution, Brain Plasticity and Visual Rehabilitation*, 6-9 June, Israel Institute for Advanced Studies, The Hebrew University of Jerusalem, Israel.

- \* Arnold, G., O'Brien, J., Spence, C., & Auvray, M. (2013). The cost of adopting and adapting tactile frames of reference. *14<sup>th</sup> Multisensory Research Forum (IMRF)*, 3-6 June, The Hebrew University of Jerusalem, Israel.
  - \* Deroy, O., Fasiello, I., Hayward, V., & Auvray, M. (2013). Audio-tactile crossmodal correspondences. *14<sup>th</sup> Multisensory Research Forum (IMRF)*, 3-6 June, The Hebrew University of Jerusalem, Israel.
  - \* O'Brien, J., Spence, C., & Auvray, M. (2013). Frames of reference at the tip of your fingers and in the palm of your hand. *14<sup>th</sup> Multisensory Research Forum (IMRF)*, 3-6 June, The Hebrew University of Jerusalem, Israel.
  - \* Girard, A., Bellik, Y., Auvray, M., & Ammi, M. (2013). Visuo-haptic tool for collaborative adjustment of selections, *8th International Workshop on Haptic and Audio Interaction Design (HAID 2013)*, 18-19 April, Daejeon, Korea (proceedings pp.40-49).
  - \* Deroy, O., & Auvray, M. (2013). A new Molyneux's problem: Sounds, shapes and arbitrary crossmodal correspondences. *Shapes 2.3: The shape of things*. 3-4 April, Rio de Janeiro, Brasil.
- Auvray, M. & Deroy, O. (2013). Audio-visual substitution. *Sensory Substitution and Augmentation Conference*, 26-28 March, British Academy, London.
- \* Girard, A., Bellik, Y., Auvray, M., & Ammi, M. (2013). Collaborative adjustment of selection areas for polygonal modelling, *IEEE 3D User Interface (3DUI'13)*, 16-17 March, Orlando, USA (proceedings pp. 135-136).
- Auvray, M. (2013). Distal attribution in sensory substitution. *Seminar From Perceptual Interaction to Extended Cognition*, 22-25 January, Compiègne, France.
- \* Girard, A., Bellik, Y., Auvray, M., & Ammi, M. (2012). Collaborative approach for dynamic adjustment of selection areas in polygonal modelling, *18<sup>th</sup> ACM symposium on Virtual Reality Software and Technology (VRST)*, 10-12 December, Toronto, Canada (proceedings pp. 205).
- Auvray, M. (2012). Substitution sensorielle: verrous et prospective. *Defisens: Audiovis*, 15-17 November, Toulouse, France.
- \* Girard, A., Ammi, M., & Auvray, M. (2012). Haptic designation strategy for collaborative molecular modeling. *IEEE International Symposium on Haptic Audio Visual Environments and Games (HAVE 2012)*, 08-09 October, Munich, Germany (proceedings, pp.13-18).
  - \* Auvray, M., & Deroy, O. (2012). The vertical integration of sensory substitution devices. *European Society for Philosophy and Psychology*, 28-31 August, London, UK.
  - \* Auvray, M., Gallace, A., & Spence, C. (2012). Crossmodal change blindness between touch, vision, and audition. *30<sup>th</sup> International Congress of Psychology (ICP)*, 22-27 July, Capetown, South Africa.
  - \* Auvray, M. (2012). Beyond vision: Integrating sensory substitution devices. *30<sup>th</sup> International Congress of Psychology (ICP)*, 22-27 July, Capetown, South Africa.
  - \* Auvray, M., Gallace, A., & Spence, C. (2012). Crossmodal change blindness between vision, touch, and audition: the influence of the task. *Association for the Scientific Study of Consciousness (ASSC 16)*, 2-6 July, Brighton, UK (poster).

- \* Auvray, M., & Deroy, O. (2012). Interpreting sensory substitution beyond the perceptual assumption: An analogy with reading. *13<sup>th</sup> International Multisensory Research Forum (IMRF)*, 19-22 June, Oxford, UK.
- Auvray, M. (2012). Spatial cognition and sensory substitution. *Mind and Perception Research Seminar*, 23 April, Glasgow, Scotland.
- Auvray, M. (2012). Interactions crossmodales. *Séminaire du Laboratoire Vision, Action, Cognition, Université Paris Descartes*, 26 March, Boulogne Billancourt, France.
- Auvray, M. (2012). A commentary of Randy Flanagan 'on sensorimotor learning'. *Workshop Perceptual Learning and Perceptual Recognition I*, 19-20 March, York University, U.K.
- Auvray, M. (2012). Interactions crossmodales: L'exemple de la cécité aux changements. *Séminaire de l'ETIS, Université de Cergy-Pontoise*, 16 March, Cergy-Pontoise, France.
- \* **Girard**, A., Ammi, M., Simard, J., & Auvray, M. (2012). Improvement of collaborative selection in 3D complex environments. *IEEE Haptics Symposium*, 4-7 March, Vancouver, Canada (proceedings, pp. 281-288).
- Auvray, M. (2012). Cognition spatiale et substitution sensorielle. *Séminaire Interactions Situées*, University Paris IV, 24 February, Paris, France.
- \* Forma, V., Hoellinger, T., Auvray, M., Roby-Brami, A., & Hanneton, S. (2011). Ears on the hand: reaching 3D audio targets. *SKILLS 2011*, 15-16 December, Montpellier, France (poster).
- Auvray, M. & Deroy, O. (2011). Understanding sensory substitution devices beyond the perceptual assumption. *Neuroscience Seminar*, 17 November, Trinity College, Hartford, USA.
- Deroy, O., & Auvray, M. (2011). Synesthesia and parasitic qualia. *Cognitive Science Seminar, CUNY*, 11 November, New York, USA.
- \* Auvray, M., Gallace, A., Yoshida, T., & Spence, C. (2011). Tactile change blindness induced by tactile and visual distractors. *12<sup>th</sup> International Multisensory Research Forum (IMRF)*, 17-20 October, Fukuoka, Japan (poster).
- Auvray, M., & Deroy, O. (2011). Understanding sensory substitution devices beyond the perceptual assumption. *Workshop Feelings, Perception, and Action*, 6-7 October, ENS, Paris.
- Auvray, M., & Deroy, O. (2011). Integrating sensory substitution devices: An analogy with reading. *Seminar of the Sensory Research Forum*, Institute of Philosophy, 6 October, London, UK.
- \* Simard, J., Ammi, M., & Auvray, M. (2011). How to improve group performances on collocated synchronous manipulation tasks. *In Proceedings of Joint Virtual Reality Conference (JVRC)*. 20-21 September, Nottingham, UK (pp. 91-94).
- \* Auvray, M. (2011). Perceiving with compensatory devices: Quasi-vision or new sensory modality? *American Philosophical Association Pacific Division*, 20-23 April, San Diego, USA.
- \* Simard, J., Ammi, M., & Auvray, M. (2010). Closely coupled collaboration for search tasks. *Virtual Reality, Software and Technology (VRST)*, 22-24 November, Hong-Kong, China (ISBN: 978-1-4503-0441-2; poster).

- \* Simard, J., Ammi, M., & Auvray, M. (2010). Study of synchronous and colocated collaboration for search tasks. *Joint Virtual Reality Conference (JVRC)*, 27-29 September, Stuttgart, Germany.
  - \* Auvray, M., Gallace, A., & Spence, C. (2010). Cécité aux changements et multimodalité. *Cognition, émotion et société, 52<sup>ème</sup> congrès de la Société Française de Psychologie*, 7-9 September, Lille, France.
  - \* Auvray, M., Gallace, A., & Spence, C. (2010). Which frame of reference is adopted when interpreting tactile letters presented on the body surface? *12<sup>th</sup> European Workshop on Imagery and Cognition (EWIC)*, 16-19 June, Helsinki, Finland.
  - \* Gallay, M., Denis, M., Parseihian, G., & Auvray, M. (2010). Egocentric and allocentric reference frames in a virtual auditory environment: Differences in navigation skills between blind and sighted individuals. *12<sup>th</sup> European Workshop on Imagery and Cognition (EWIC)*, 16-19 June, Helsinki, Finland.
- Auvray, M. (2010). Perceptual interactions in a minimalist tactile environment. *Seminar of the Department of Psychiatry and Psychotherapy*, University of Cologne, 15 April, Cologne, Germany.
- Deroy, O., & Auvray, M. (2010). A new look on sensory extensions. *Workshop on Sensory Substitution, Synaesthesia, Sensation, and Perception*, Institute of Philosophy, 31 March, London, UK.
- Auvray, M. (2010). Cross-modal change blindness. *Seminar of the Sensory Research Forum*, Institute of Philosophy, 30 March, London, UK.
- Auvray, M. (2010). Percevoir avec des dispositifs de compensation: de la substitution sensorielle à l'extension sensorimotrice. *Phiteco*, 18-24 January, Compiègne, France.
- \*Auvray, M., Gallace, A., & Spence, C. (2009). What perspective do people take when interpreting tactile letters presented on their bodies? *10<sup>th</sup> International Multisensory Research Forum (IMRF)*, 29 July - 2 August, New-York, USA (poster).
  - \* Auvray, M., Gallace, A., & Spence, C. (2009). Short term memory for tactile stimuli presented on the fingertips and across the body surface. *10<sup>th</sup> International Multisensory Research Forum (IMRF)*, 29 July - 2 August, New-York, USA (poster).
  - \* Auvray, M., Hoellinger, T., Roby-Brami, A., & Hanneton, S. (2009). Perceptual weight judgments when viewing own and other's movements under minimalist conditions of visual presentation. *10<sup>th</sup> International Multisensory Research Forum (IMRF)*, 29 July - 2 August, New-York, USA (poster).
- Auvray, M., & Myin, E. (2009). Perception with compensatory devices: From sensory substitution to sensorimotor extension. *The Senses Research Workshop*, 19-21 June, Paris, France.
- Auvray, M. (2009). NAVIG: Navigation Assisted by artificial VIsion and GNSS. *International Workshop on Mental Imagery, Spatial Cognition, and Language (MICL 2009)*, in honour of Michel Denis, 12 June, Padova, Italy.
- Auvray, M. (2009). Perceptual interactions in a minimalist environment. *The Third Cargese Consciousness Summer School, Consciousness of the Self - Consciousness of the Others*, 25 May - 4 June, Cargese, France.
- Auvray, M. (2009). Replacing one sense by another: Sensory substitution and the classification of our sensory modalities. *Key Issues in Sensory Augmentation Workshop*, 26-27 March, Brighton, UK.

- \* Auvray, M., Hoellinger, T., & Hanneton, S. (2008). Sensory substitution and the taxonomy of our sensory modalities. *9<sup>th</sup> Annual Meeting of International Multisensory Research Forum (IMRF)*, 16-19 July, Hamburg, Germany.
- \* Hoellinger, T., Auvray, M., Roby-Brami, A. & Hanneton, S. (2008). Localisation tasks with a three-dimensional audio-motor coupling based on an electromagnetic motion capture device. *9<sup>th</sup> Annual Meeting of International Multisensory Research Forum (IMRF)*, 16-19 July, Hamburg, Germany (poster).
- Auvray, M. (2008). Remplacer un sens par un autre: les dispositifs de substitution sensorielle. *9<sup>ème</sup> Journée de la Vision, SFOP*, 13 April, Paris, France.
- Auvray, M. (2008). The multisensory nature of spatial attention: Crossmodal change blindness between vision, touch, and audition. *Séminaire du Laboratoire de Neurophysique et Physiologie*, 28 January, Paris, France.
- Auvray, M. (2007). Perception spatiale et multisensorialité: Espaces réels et espaces virtuels. *Séminaire du LIMSI: CHM et VENISE*, 04 December, Orsay, France.
- Gallace, A., Auvray, M., Hartcher-O'Brien, J., Tan, H. Z., & Spence, C. (2007). The effect of visual, auditory, and tactile distractors on people's awareness of tactile change. *Body Representation Workshop*, 8-10 October, Rovereto, Italy.
- \* Auvray, M., Myin, E., & Spence, C. (2007). Crossmodal change blindness and the multisensory nature of spatial attention. *International Conference on Perception and Action (ICPA 14)*, 1-6 July, Yokohama, Japan.
- \* Lenay, C., Stewart, J., Ali Ammar, A., & Auvray, M. (2007). Direct perception of an intentional subject: A minimalist approach. *International Conference on Perception and Action (ICPA 14)*, 1-6 July, Yokohama, Japan (poster).
- \* Auvray, M. (2007). Enaction and sensory substitution. *Philosophy of Psychology, Neuroscience, and Psychology*, 25 March, Bristol, UK.
- Auvray, M. (2007). Cognition spatiale et appropriation d'outils. *Séminaire du LIMSI: CHM et VENISE*, 06 March, Orsay, France.
- Lenay, C., Auvray, M., Sebbah, F.-D., & Stewart J. (2006). Perception of an intentional subject: An enactive approach, in *Proceedings of the Third International Conference on Enactive Interface*, 20-21 November, Montpellier, France, pp. 37-38.
- Auvray, M. (2006). *Perception, interaction, and mutual recognition in sensory substitution. Action in Mind, The 6<sup>th</sup> Annual Conference on the Philosophy of Consciousness of the University of Antwerp*, 21 October, Antwerp, Belgium.
- \* Lenay, C., Auvray, M., Thouvenin, I., & Sebbah, F. D. (2006). Enactive approach to the recognition of an intentional subject. *Situated Cognition: Perspectives From Phenomenology and Science*, 18-20 August, Durham, UK.
- \* Auvray, M., Lenay, C., & Stewart, J. (2006). The recognition of others in a simulated environment, the case of minimalist devices. *Association for the Scientific Study of Consciousness 10*, 23-26 June, Oxford, UK (poster).
- \* Lenay, C., Auvray, M., & Sebbah, F. D. (2006). Seeing oneself: Experimental studies of perceptive interactions. *Association for the Scientific Study of Consciousness 10*, 23-26 June, Oxford, UK (poster).

- \* Auvray, M., Gallace, A., & Spence, C. (2006). Crossmodal change blindness between vision and touch. *7<sup>th</sup> Annual Meeting of International Multisensory Research Forum (IMRF)*, 18-21 June, Dublin, Ireland (poster).
- \* Auvray, M., Gallace, A., & Spence, C. (2006). Crossmodal change blindness. *EPS Birmingham Meeting*, 10-12 April, Birmingham, UK.
- Auvray, M. (2005). A commentary of Zenon Pylyshyn: “*Spatial representation in the mind/brain: Do we need a global topographical map?*” *Frames of Reference in Spatial Cognition: Diversity, Flexibility and Adaptability*, 17-19 November, Paris, France.
- \* Lenay, C., Auvray, M., Thouvenin, I., & Sebbah, F. D. (2005). Perception of avatars in a space of tactile interaction. *2nd Intuition International Workshop “VR/VE and Industry: Challenges and Opportunities”* (ISBN 960-254-658-1), 24-25 November, Paris, France.
- Auvray, M. (2005). Spatial perception and sensory substitution. *Body Representation Workshop*, 30 September, Oxford, UK (Poster).
- \* Auvray, M., Lenay, C., Hanneton, S., Gapenne, O., & Stewart, J. (2005). Spatial perception and sensory substitution. *International Conference on Perception and Action*, 5-10 July, Monterey, USA.
- \* Gapenne, O., Rovira, K., Lenay, C., Stewart, J., & Auvray, M. (2005). Is form perception necessarily tied to specific sensory feedback? *International Conference on Perception and Action*, 5-10 July, Monterey, USA.
- Auvray, M. (2005). Perception spatiale et substitution sensorielle. *Journées GDR-STIC Santé, Espace et sa Perception: Application à la Conception des Aides Technologiques aux Déplacements de Déficients Visuels*, 10 March, Paris, France.
- Auvray, M. (2005). Appropriation d'une prothèse perceptive et constitution de nouveaux espaces de perception et d'action. *Journée Atelier Rescif, Indétermination du Champ Visuel, Anticipations Catégorielles et Thématization*, 1 February, Paris, France.
- \* Auvray, M., Hanneton, S., Lenay, C., & O'Regan, J. K. (2004). Exteriorisation in sensory substitution. *First Joint Conference of SPP & ESPP*, 3-6 July, Barcelona, Spain.
- \* Auvray, M., Hanneton, S., O'Regan, J. K., & Lenay, C. (2004). Distal attribution in sensory substitution. *Association for the Scientific Study of Consciousness 8*, 25-28 June, Antwerp, Belgium.
- Auvray, M., Lenay, C., O'Regan, J. K., & Hanneton, S. (2004). Action et Immersion. *Séminaire Savoir ce que l'on Fait*, 27-31 January, Compiègne, France.
- \* Auvray, M., Hanneton, S., & O'Regan, J. K. (2003). Localisation, form recognition and emergence of sensation in sensory substitution. *ESCOP*, 17-20 September, Granada, Spain.
- \* Auvray, M., Hanneton, S., & O'Regan, J. K. (2003). Localisation and form recognition in sensory substitution. *ECVP*, 1-5 September, Paris, France. *Perception*, 32, supplement, 94 (poster).
- Auvray, M., O'Regan, J. K., & Hanneton, S. (2003). Repenser la distinction entre nos modalités sensorielles via un dispositif de suppléance perceptive visuo-auditif. *Colloque des Jeunes Chercheurs en Sciences Cognitives*, 18-20 June, Paris, France.
- Khatchatourov, A., & Auvray, M. (2003). L'outil modifie-t'il la perception ou la rend-il possible ? *Colloque des Jeunes Chercheurs en Sciences Cognitives*, 18-20 June, Paris, France.

- Auvray, M., O'Regan, J. K., & Hanne-ton, S. (2003). Apprendre à utiliser un dispositif de substitution sensorielle: les différentes étapes. *Séminaire Espaces d'Action, Espaces de Perception*, 21-29 January, Compiègne, France.
- Auvray, M., Hanne-ton, S., & O'Regan, J. K. (2002). Stages in the emergence of sensation in sensory substitution. *Third Workshop on Multimodal Interaction in Perception*, 28-29 October, Paris, France.
- Auvray, M., Philipona, D., & O'Regan, J. K. (2001). Tentative d'induire la notion d'espace dans un environnement virtuel simple. *Atelier Perception et Constitution de l'Espace, Journées du RESCIF*, 27-28 October, Paris, France.
- Auvray, M., Philipona, D., & O'Regan, J. K. (2001). Notion of space and form recognition in a simulated environment. *Atelier Genèse de la Perception et de la Notion d'Espace chez le Robot et l'Homme*, 18-19 October, Paris, France.
- O'Regan, J. K., Clark, J. J., Nadal, J-P., Auvray, M., Bompas, A., & Philipona, D. (2001). Implications of a "sensorimotor" theory of visual phenomenality. *International Workshop: Pathways to Abstraction: From Experience to Concept*, 10-13 September, Gif-sur-Yvette, France.
- \* Auvray, M., & O'Regan, J. K. (2001). Influence of semantic factors on blindness to progressive changes in visual scenes. *ECVP*, 26-30 August, Kusadasi, Turkey. *Perception*, 30, supplement, 44 (poster).

## Annex

### Selection of articles

1. Pesnot-Lerousseau, J., Arnold, G., & Auvray, M. (submitted). Visualizing sounds: training-induced plasticity with a visual-to-auditory conversion device.

*As this article is in the submitted stage, it is not included for now in the on-line version of the HDR*

2. Arnold, G., & Auvray, M. (2014). Perceptual learning: Tactile letter recognition transfers across body surfaces. *Multisensory Research*, 27, 71-90.
3. Arnold, G., & Auvray, M. (2017). Tactile recognition of visual stimuli: Specificity versus generalization of perceptual learning. *Vision Research*.
4. Deroy, O., Fasiello, I., Hayward, V., & Auvray, M. (2016). Differentiated audio-tactile correspondences in sighted and blind individuals. *Journal of Experimental Psychology: Human Perception and Performance*, 42.
5. Arnold, G., Spence, C., & Auvray, M. (2016). Taking someone else's spatial perspective: Natural stance or effortful decentring? *Cognition*, 148, 27-33.
6. O'Brien, J., & Auvray, M. (2016). Cognition overrides orientation dependence in tactile viewpoint selection. *Experimental Brain Research*, 234, 1885-1892.
7. O'Brien, J., & Auvray, M. (2014). The process of distal attribution illuminated through studies of sensory substitution. *Multisensory Research*, 27, 421-441.



# Perceptual Learning: Tactile Letter Recognition Transfers Across Body Surfaces

Gabriel Arnold<sup>1,2,\*</sup> and Malika Auvray<sup>2</sup>

<sup>1</sup> LIMSI-CNRS, UPR 3251, rue John von Neumann, 91400 Orsay, France

<sup>2</sup> Institut Jean Nicod, UMR 8129, Ecole Normale Supérieure, Paris, France

Received 27 September 2013; accepted 15 February 2014

---

## Abstract

Visual-to-tactile sensory substitution devices are designed to assist visually impaired people by converting visual stimuli into tactile stimuli. The important claim has been made that, after training with these devices, the tactile stimuli can be moved from one body surface to another without any decrease in performance. This claim, although recurrent, has never been empirically investigated. Moreover, studies in the field of tactile perceptual learning suggest that performance improvement transfers only to body surfaces that are closely represented in the somatosensory cortex, i.e. adjacent or homologous contralateral body surfaces. These studies have however mainly used discrimination tasks of stimuli varying along only one feature (e.g., orientation of gratings) whereas, in sensory substitution, tactile information consists of more complex stimuli. The present study investigated the extent to which there is a transfer of tactile letter learning. Participants first underwent a baseline session in which the letters were presented on their belly, thigh, and shin. They were subsequently trained on only one of these body surfaces, and then re-tested on all of them, as a post-training session. The results revealed that performance improvement was the same for both the trained and the untrained surfaces. Moreover, this transfer of perceptual learning was equivalent for adjacent and non-adjacent body surfaces, suggesting that tactile learning transfer occurs independently of the distance on the body. A control study consisting of the same baseline and post-training sessions, without training in between, revealed weaker improvement between the two sessions. The obtained results support the claim that training with sensory substitution devices results in a relative independence from the stimulated body surface.

## Keywords

Perceptual learning, touch, sensory substitution, generalization, tactile letter recognition

---

This article is part of the 14th IMRF collection, guest edited by A. Amedi, D. I. Shore and S. Levy-Tzedek.

\* To whom correspondence should be addressed. E-mail: gabriel.arnold@limsi.fr

## 1. Introduction

Sensory substitution devices convert stimuli that are normally accessed through one sensory modality (e.g., vision) into stimuli accessible through another sensory modality (e.g., touch or audition). Among them, visual-to-tactile and visual-to-auditory conversion systems were designed to assist blind and visually impaired people by converting visual stimuli into tactile or auditory stimuli.

The first sensory substitution device was the tactile-vision sensory substitution (TVSS) developed by Bach-y-Rita and colleagues (Bach-y-Rita *et al.*, 1969; see also Bach-y-Rita and Kercel, 2003, for a review). The TVSS comprises a camera, a translation system, and a matrix of tactile stimulators (that provides electrical or vibratory stimulation) placed on a body surface such as the back or the tongue (see Bach-y-Rita and Kercel, 2003, for a description). The principles of sensory substitution have been extended to the conversion of visual images into sounds. This led to the design of visual-to-auditory devices such as the vOICe (Meijer, 1992), the prosthesis for substitution of vision by audition (PVSA, see Capelle *et al.*, 1998), and the Vibe (Hanneton *et al.*, 2010).

Since their invention, a variety of devices have been designed and tested across a variety of tasks (see Auvray and Myin, 2009; Bubic *et al.*, 2010; Deroy and Auvray, 2012, for reviews). In particular, studies have revealed that these devices allow their users to perform localization tasks (Jansson, 1983; Levy-Tzedek *et al.*, 2012; Renier *et al.*, 2005) and simple as well as complex shape recognition (Arno *et al.*, 2001; Auvray *et al.*, 2007; Pollok *et al.*, 2005; Sampaio *et al.*, 2001). A defining feature of perception with sensory substitution devices is that extensive training is required. Most users are able, in one or two hours, to explore their environment and to approximate objects' position and shape. However, the training required to obtain more precise object localization and recognition behaviour is estimated at around eight hours with visual-to-tactile devices (Kaczmarek and Haase, 2003) and 10 to 15 with visual-to-auditory devices (Auvray *et al.*, 2007). Perceptual acuity (measured with the Snellen tumbling E test) has also been reported to increase with training, reaching for instance 20/430 after nine hours of training with the Tongue Display Unit (Sampaio *et al.*, 2001; see also Haigh *et al.*, 2013, for a similar study with the vOICe).

One important claim of Bach-y-Rita is that, once trained, TVSS-users no longer feel the stimulation on their skin, where it occurs, but they directly attribute the stimulation as resulting from an external object, i.e. located at a distance (Bach-y-Rita, 1995, p. 181; Bach-y-Rita and Kercel, 2003, p. 543). Furthermore, Bach-y-Rita asserted that, consequent to this externalization process, the tactile stimulator array can be moved from one body surface to

another (e.g., from the back to the abdomen or to the forehead), without loss in spatial localization abilities or other perceptual capacities (see Note 1) (such as depth estimates). However, this claim was based solely on users' verbal reports and it has never been experimentally investigated.

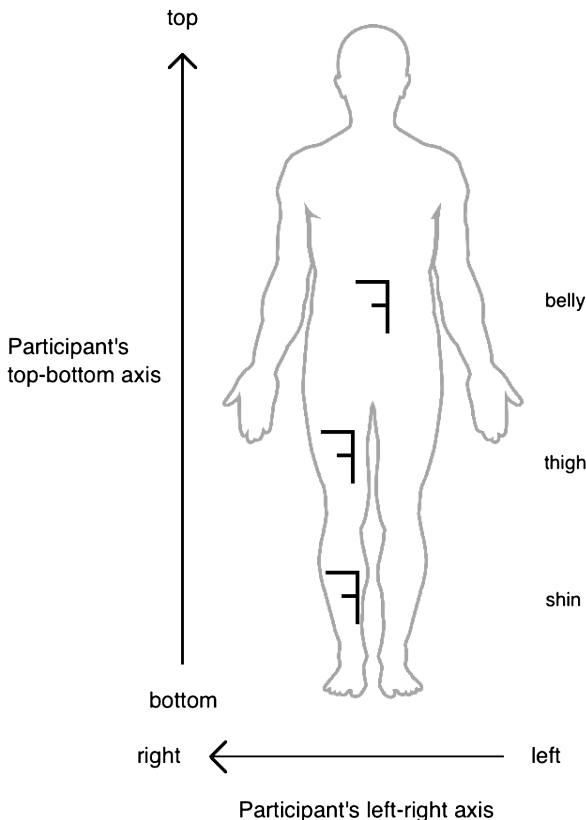
One way to investigate if perceptual abilities in trained users of visual-to-tactile sensory substitution devices depend on the stimulated surface or not, is to study the extent to which there is a transfer of tactile learning across body surfaces. If indeed what is learned on a given body surface can be transferred to another one without loss in performance, this would provide reasons to believe that Bach-y-Rita's claim of independence from the stimulated body surface is correct. Perceptual learning is defined as performance improvement in perceptual tasks resulting from training (Fahle and Poggio, 2002; Gibson, 1969). In tactile perceptual learning, the question of the specificity (i.e., learning is body-surface dependent) *versus* generalization (i.e., there is a transfer of performance improvement from a trained to an untrained body surface) is still debated. Some studies reported a transfer to untrained body surfaces (Harrar *et al.*, 2014; Harris *et al.*, 2001; Kaas *et al.*, 2013; Nagarajan *et al.*, 1998; Sathian and Zangaladze, 1997; Spengler *et al.*, 1997) whereas other studies did not find evidence of such transfer (Dinse *et al.*, 2006; Godde *et al.*, 2000). In addition, when transfer has been reported, it was restricted to adjacent and homologous contralateral body surfaces only (e.g., adjacent fingers of the same hand and same fingers of the untrained hand — Harrar *et al.*, 2014; Harris *et al.*, 2001). In other words, transfer to non-adjacent and non-homologous body surfaces has never been reported. The extent to which there is a transfer of tactile learning therefore appears to be related to the topographic organization of the somatosensory cortex.

There is thus an apparent contradiction between this reported limit to the transfer of tactile learning (i.e., restricted to adjacent and homologous body surfaces) and the post-training independence from the stimulated body surface (i.e., occurring even in non-adjacent surfaces such as the back and the abdomen), as claimed by Bach-y-Rita. One hypothesis to account for such a discrepancy is that the discrimination tasks used in tactile learning studies involved low-level stimuli that mainly vary along only one feature (e.g., gratings with different orientations), whereas reports in the field of sensory substitution were based on more complex objects consisting of combinations of different features. Discrimination of low-level stimuli may involve somatosensory areas with a topographic representation of body surfaces. In these areas, receptive fields are narrow and overlap only with receptive fields of adjacent or contralateral homologous body surfaces (for instance, see Iwamura *et al.*, 1994, for the topographic organization of the hands). On the other hand, sensory substitution devices are mainly used to convey tactile information resulting from higher-level stimuli such as objects, faces, and scenes. These high-level stim-

uli involve less topographically organized areas in which receptive fields are large and not restricted to a specific body surface. Moreover, according to the reverse hierarchy theory (Ahissar and Hochstein, 2004), perceptual learning begins at higher cortical levels and continues at lower levels of processing dependent on the feature learnt. As a result, tactile learning of high-level stimuli should transfer more easily to different body surfaces than low-level stimulus features.

The study reported here aims at investigating the extent to which there is a transfer of tactile learning in the recognition of high-level symbols (i.e., made of the combinations of several features rather than consisting of single features). To do so, a tactile letter recognition task was used. Letters were drawn on the participants' body by means of sequential vibrotactile stimulations (see Yanagida *et al.*, 2004, for a similar method). In order to evaluate learning effects rather than mere repeated exposure effects, we measured recognition performance before and after training, as in Harrar *et al.*'s (2014) study. The participants first completed a baseline session in which they had to recognize tactile letters drawn on three different body surfaces: the belly, the front of the right thigh, and the right shin (see Fig. 1 below). They then underwent a training session in which the letters had to be recognized on only one of the three above-mentioned body surfaces. Finally, the participants performed a post-training session on all three body surfaces. The amount of tactile learning was evaluated by computing performance improvement (both accuracy and response times) between the baseline and the post-training sessions. If there is a transfer of tactile learning from a trained to an untrained body surface, then the amount of performance improvement should be similar for trained and untrained surfaces. On the other hand, if tactile learning is specific to a given body surface, then performance improvement should be greater for trained than for untrained body surfaces. In addition, a control group of participants performed the baseline and post-training sessions, but without the training session in between. For these participants, performance is expected to be either similar between the two sessions, or to improve in the second one but less than for the participants in the trained group.

It should be mentioned that the belly, the thigh and the shin are represented in adjacent areas of the somatosensory cortex (Merzenich *et al.*, 1978; Nakamura *et al.*, 1998). Thus, performance improvement, computed as a function of the trained body surface, allows behaviourally evaluating the topographic organization of areas in which tactile learning of high-level symbols occurs. If tactile learning occurs in non-topographic areas, the amount of improvement should not depend on the trained body surface. On the other hand, if tactile learning occurs in topographic areas, the amount of improvement should be greater for two adjacent body surfaces (e.g., the belly and the thigh) than for two non-adjacent ones (e.g., the belly and the shin).



**Figure 1.** Example of the letter F drawn on the three body surfaces (belly, thigh, and shin). The letters' left-right and top-bottom axes were always congruent to the participants' left-right and top-bottom axes.

## 2. Methods

### 2.1. Participants

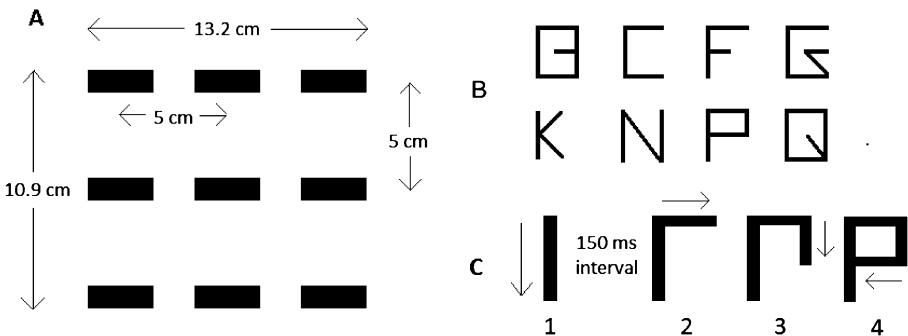
Forty-five participants completed the experiment (26 females and 19 males; mean age = 25.7 years, range = 18–47 years). Thirty participants were included in the trained group and 15 in the control group. All the participants were naïve to the purpose of the experiment. They provided their informed consent and received payment for their participation. The experiment took approximately one hour and a half to complete (one hour for the control group) and was performed in accordance with the ethical standards laid down in the Declaration of Helsinki (1991).

It should be mentioned that there was a selection process used to determine if participants were included in the study. Given that tactile letters drawn on the skin can be interpreted within different reference frames (see Parsons and

Shimojo, 1987; Sekiyama, 1991), each participant's preferred reference frame was identified prior to the start of the experiment. This was done by means of a task requiring the participants to recognize the tactile letters b, d, p, and q. Recognizing these ambiguous asymmetrical letters requires assigning the top-bottom and left-right axes to the letters, and the reference frame that was adopted can be determined from how the stimulus is interpreted. Only the participants that freely assigned the top-bottom and left-right axes congruently to their body's axes were selected for participation in the study. The letters were then drawn consistently to these body-congruent assignments during the rest of the experiment (see Fig. 1). The reason for this choice is that — as the aim of the study was not to investigate the influence of reference frame on tactile learning — we wanted to avoid any interference due to differences in the reference frames that are freely adopted. We consequently chose the most frequently adopted one among the four possible reference frames. This reference frame with body-congruent assignments of axes was adopted by 52% of participants across all the experiments that were conducted in our laboratory ( $N = 244$ ). Note that 28% of participants reversed the top-bottom axis, while the left-right axis was consistent with the body. A further 20% of participants reversed the left-right axis, while the top-bottom axis was consistent with the body. No participants reversed both the vertical and horizontal axes.

## 2.2. Apparatus

The tactile stimuli were presented by means of nine rectangular vibrators (Haptuator Mark II, Tactile Labs, Montreal, Canada) arranged in a three-by-three array with a centre-to-centre spacing of 5 cm (see Fig. 2A). The surface area of each vibrator was 0.9 cm vertically by 3.2 cm horizontally. The vibrator array was placed on the participants' body surface, above their clothes, by means of an elastic belt. When the participants were tested on the belly, the vibrator array was placed symmetrically to their body mid-sagittal line



**Figure 2.** (A) Schematic figure illustrating the  $3 \times 3$  array of rectangular vibrators. (B) The 8 letters used as tactile stimuli. (C) The sequence of vibrations chosen for drawing the letter P.

(see Fig. 1). In this case, the three lower factors were located above the waist. When they were tested on the front of the thigh and on the shin, the vibrator array was both horizontally and vertically centred. The position of the vibrator array on each body surface was exactly the same between the baseline and the post-training sessions; this was achieved by means of stickers indicating the positions of the vibrators.

A nine-channel amplifier drove each vibrator independently at a 250 Hz frequency. The intensity of each vibrator was individually selected with an adjustment method at the beginning of each session. Specifically, the participants were instructed to adjust the vibration intensity until each vibrotactile stimulus could be perceived clearly and with the same felt intensity. The activation of the vibrators was controlled through a PC running custom software written in MATLAB R2008a. Instructions and feedback were presented on a 23-inch screen with a 1920 × 1080 resolution. Participants wore noise-reducing headphones with a noise reduction rating of 30 dB, in order to mask any sounds made by the vibrators.

### 2.3. Stimuli

The eight upper-case letters B, C, F, G, K, N, P, and Q were drawn on the participants' body surfaces (see Fig. 2B). These letters were selected on the basis of previous experiments we conducted, in which they were sufficiently well recognized (above 87% accuracy) and induced little confusion. The order in which the vibrators were successively activated was chosen to approximate the gesture of manual letter drawing (see Fig. 2C). This sequential mode of presentation was chosen over a simultaneous mode because it allows better performance in letter recognition (Yanagida *et al.*, 2004; see also Loomis, 1974). In order to choose the sequence of each letter, 14 persons, who did not participate in the main experiment, were asked to draw the letters on a paper sheet. From this we calculated the most frequent way of writing each of the letters. For instance, the letter P is the most frequently drawn by an initial vertically descending stroke followed by a loop beginning from the top of the vertical stroke. Each vibrator composing the sequence was activated for 250 ms. There was no interval between two consecutive vibrations composing one stroke of the letter (e.g., the three vibrators composing the vertical stroke of the letter P); nor between two consecutive strokes that were spatiotemporally continuous (e.g., the three strokes of the loop of the letter P). However there was a 150-ms interval between two consecutive strokes that were spatiotemporally discontinuous (e.g., the end of the vertical descending stroke and the beginning of the loop of the letter P). This interval avoided possible errors of interpretation caused by an irrelevant grouping of vibrations. Participants were instructed that these temporal intervals corresponded to a

spatial discontinuity in manual letter drawing. The mean duration of letters was 2338 ms.

## 2.4. Procedure

### 2.4.1. Baseline and Post-Training Sessions

The participants were required to stand in front of the computer screen. On each trial, one of the eight letters was drawn on the participant's body surface. The participants were instructed to report which letter was drawn by pressing the corresponding key on the computer keyboard with the index finger of their dominant hand. Stickers were used to represent the eight letters on two lines of four adjacent keys. Participants were required to respond as accurately and as rapidly as possible. Accuracy was emphasized over speed of response. Given that posture has an influence on how tactile symbols are perceived (see Parsons and Shimojo, 1987; Sekiyama, 1991), we made sure that the participant's posture was consistent across sessions. As they had to look at the screen when visual feedback was provided to them, the participants were instructed to look at the empty screen during the conditions in which there was no feedback. However, note that, in all the conditions, the participants were allowed to look at the keyboard when giving their answers. The participants were able to give their responses at any time from the onset of the first vibration and up to 3000 ms after the end of the last vibration (at which point the trial was terminated). At the end of each trial, there was an interval of 3000 ms before the beginning of the next trial. No feedback was given regarding the correctness of the participants' responses.

The baseline and the post-training sessions were composed of three blocks each: one block for each body surface. The six possible orders of these three blocks were counterbalanced across participants. However, the order remained the same for each participant during the baseline and the post-training sessions. The three body surfaces were thus each separated by the same amount of letter expositions (i.e., five blocks of trials) between the baseline and the post-training sessions. Each block was composed of 48 randomized trials corresponding to six presentations of each letter.

In the baseline session, before each of the experimental blocks, the participants were given eight practice trials, corresponding to one presentation of each letter. In this practice block, feedback was given after each response: the correct answer was visually displayed on the computer screen once participants responded.

### 2.4.2. Training Session

The participants in the control group completed the post-training session just after the baseline session. Only the participants of the trained group were trained in between the two sessions. During this training session, the participants were trained on only one of the three body surfaces. This trained surface

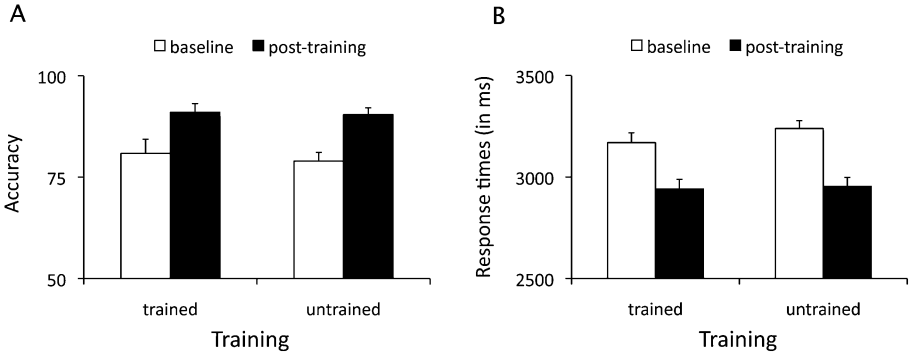


was always the surface that was tested in the second position during the baseline and the post-training sessions. Thus, one third of the participants from the trained group were trained on each of the possible body surfaces. The training session consisted of three consecutive blocks of 48 trials. Given that a trial-by-trial feedback has been shown to reinforce learning (Herzog and Fahle, 1998), in this session, the participants were given visual feedback on the correctness of their answers after each trial.

### 3. Results

Trials in which participants failed to make a response before the trial was terminated (less than 0.9% of trials overall) were not included in the data analyses. The accuracy of letter recognition was measured by the percentage of correct responses in each condition. The mean response times (RTs) were also computed. However, RTs corresponding to errors were excluded from the RT analysis. The first analysis presented here evaluated training effects by measuring performance improvement across sessions for the trained and control groups. The second analysis evaluated the transfer of learning by comparing training effects in trained and untrained body surfaces for the trained group only. The third one evaluated the possible topographic organization of perceptual transfer by measuring performance improvement as a function of the different body surfaces in the trained group. For all analyses, an alpha level of 0.05 was used.

In order to evaluate the effect of training on performance, ANOVAs were conducted both on accuracy and on RTs with Group (trained, control) as a between-participant factor and Session (baseline, post-training) as a within-participant factor. Because the sample sizes were unequal, type III sum of squares were used (Shaw and Mitchell-Olds, 1993). With respect to accuracy, there was a significant effect of Session ( $F(1, 43) = 45.42$ ;  $p < 0.001$ ). Accuracy was greater in the post-training (86.51%,  $SD = 11.72$ ) than in the baseline (77.49%,  $SD = 12.67$ ) sessions, thus showing performance improvement. Importantly, there was a significant interaction between Group and Session ( $F(1, 43) = 6.35$ ;  $p < 0.05$ ), showing that performance improvement was greater in the trained (increase of 11.02 points of percentage,  $SD = 8.34$ ) than in the control group (5.02 points of percentage,  $SD = 5.47$ ). As a consequence, global accuracy (i.e., across the two sessions) was also significantly greater ( $F(1, 43) = 7.27$ ;  $p < 0.01$ ) for the trained (85.07%,  $SD = 12.69$ ) than for the control group (75.87%,  $SD = 11.35$ ). Specific comparisons showed that the group difference was significant only for the post-training session ( $F(1, 43) = 14.06$ ;  $p < 0.001$ ); not for the baseline session ( $F(1, 43) = 2.48$ ;  $p = 0.12$ ). In other words, prior to training, accuracy in letter recognition was similar for the two groups. However, after training, recognition improved



**Figure 3.** Participants' (A) accuracy and (B) response times of the trained group for the trained and untrained body surfaces obtained during the baseline session (in white) and the post-training session (in black). Error bars represent the standard errors of the means.

differentially as the percentage of correct responses was significantly higher for the trained group than those obtained for the control group. With respect to RTs, there was a significant effect of Session ( $F(1, 43) = 102.24$ ;  $p < 0.001$ ). RTs were shorter in the post-training (3067 ms,  $SD = 282$ ) than in the baseline (3319 ms,  $SD = 259$ ) sessions. RTs were also significantly shorter ( $F(1, 43) = 21.62$ ;  $p < 0.001$ ) for the trained (3088 ms,  $SD = 268$ ) than for the control group (3403 ms,  $SD = 238$ ). The interaction between Group and Session was not significant ( $F(1, 43) = 1.88$ ;  $p = 0.18$ ).

In order to determine if there was a transfer of tactile learning, ANOVAs were conducted with Session (baseline, post-training) and Training (trained, untrained) as within-participant factors for the trained group only (see Fig. 3A and 3B). As with the previously reported ANOVA, accuracy was significantly greater in the post-training than in the baseline sessions ( $F(1, 29) = 42.70$ ;  $p < 0.001$ ). There was no effect of Training ( $F(1, 29) < 1$ ). More importantly, there was no significant interaction between Session and Training ( $F(1, 29) < 1$ ); which shows that performance improved between the baseline and post-training sessions for both trained (increase of 10.13 points of percentage,  $SD = 12.32$ ) and untrained (11.47 points of percentage,  $SD = 7.79$ ) body surfaces. The analysis conducted on RTs yielded the same pattern of results, with a significant decrease in RTs from the baseline to the post-training sessions ( $F(1, 29) = 85.10$ ;  $p < 0.001$ ). There was a significant effect of Training ( $F(1, 29) = 4.32$ ;  $p < 0.05$ ) but no significant interaction between Session and Training ( $F(1, 29) = 3.37$ ;  $p = 0.08$ ). Note that although the interaction approached significance, the RT analysis did not show a greater training effect for trained than untrained body surfaces. The decrease in RTs was 228 ms ( $SD = 173$ ) for trained and 285 ms ( $SD = 176$ ) for untrained body surfaces. Thus, for both accuracy and RTs, the performance improvement

obtained with training was similar for trained and untrained body surfaces, showing that perceptual learning transfers to untrained surfaces.

Finally, in order to evaluate the possible influence of topographic organization on tactile perceptual learning, ANOVAs were conducted with the Trained surface (belly, thigh, shin) as a between-participant factor, and with Session (baseline, post-training) and Stimulated surface (belly, thigh, shin) as within-participant factors for the trained group only (see Tables 1 and 2). As with previous analyses, the performance improvement from the baseline to the post-training sessions was significant for both accuracy ( $F(1, 29) = 49.77$ ;  $p < 0.001$ ) and RTs ( $F(1, 29) = 87.26$ ;  $p < 0.001$ ). There was also a significant effect of the Stimulated surface for both accuracy ( $F(2, 58) = 11.92$ ;  $p < 0.001$ ) and RTs ( $F(2, 58) = 5.13$ ;  $p < 0.01$ ). Tukey's HSD tests showed that performance was significantly better ( $p < 0.01$ , for accuracy;  $p < 0.05$ , for RTs) for the belly (89.72%, SD = 10.85; 3028 ms, SD = 279) than for the thigh (83.71%, SD = 14.28; 3114 ms, SD = 302) and significantly better ( $p < 0.001$ , for accuracy;  $p < 0.05$ , for RTs) for the belly than for the shin (81.78%, SD = 16.91; 3121 ms, SD = 310), but there was no significant difference between the thigh and the shin. More importantly, there was neither a significant interaction between Trained surface and Session ( $F(2, 27) < 1$ , for both accuracy and RTs) nor any significant interaction between Trained surface, Stimulated surface and Session ( $F(4, 54) < 1$ , for both accuracy and RTs), indicating that performance improvement did not vary with the distance between the trained and stimulated surfaces. Performance was significantly better ( $p < 0.05$  for all comparisons) in the post-training than in the baseline sessions for all the combinations of trained and stimulated surfaces, i.e., including adjacent and non-adjacent body surfaces (for both accuracy and RTs). There was no other significant effect or interaction.

The confusion matrices of the trained participants' responses (see Table 3) indicate that before training, the participants made the same types of confusions for the three stimulated surfaces. In particular, the letters K and F were reciprocally confused and the letter G was frequently mistaken for the letter Q. In addition, the letter N was frequently mistaken for the letter P when it was displayed on the shin but not when it was displayed on the belly and the thigh. Overall, the same types of letter confusion were observed after training, with however less errors than before training.

#### 4. Discussion

The study reported here aimed at investigating the extent to which a performance improvement in tactile letter recognition transfers across body surfaces. Three main results emerged from this study. First, training improved both participants' accuracy and response latency. In addition, participants performing

**Table 1.**

Mean accuracy (in percentages) during the baseline and post-training sessions and the corresponding training effects, as a function of stimulated surface (belly, thigh, shin), and trained surface (belly, thigh, shin, untrained). Note that the participants were tested on all three surfaces but trained on only one of them. Standard deviations of the mean are in brackets

Surface trained	Baseline			Post-training			Training effect		
	Belly	Thigh	Shin	Belly	Thigh	Shin	Belly	Thigh	Shin
Belly	87.9 (8.8)	77.5 (11.9)	71.7 (14.0)	95.2 (7.8)	89.6 (6.7)	87.3 (13.4)	7.3 (9.9)	12.1 (10.7)	15.6 (8.9)
Thigh	81.5 (14.7)	74.8 (22.6)	74.6 (16.9)	92.3 (8.6)	87.5 (11.7)	87.1 (13.2)	10.8 (11.0)	12.7 (13.1)	12.5 (12.9)
Shin	87.1 (10.7)	81.3 (13.4)	79.8 (22.3)	94.4 (8.7)	91.7 (7.7)	90.2 (14.9)	7.3 (8.0)	10.4 (11.2)	10.4 (14.2)
Untrained (control)	84.6 (9.5)	69.4 (16.0)	66.0 (16.3)	86.2 (11.6)	73.8 (15.7)	75.2 (13.6)	1.6 (6.1)	4.4 (12.1)	9.2 (12.1)

**Table 2.**

Mean response times (in ms) during the baseline and post-training sessions and the corresponding training effects, as a function stimulated surface (belly, thigh, shin), and trained surface (belly, thigh, shin, untrained). Standard deviations of the mean are in brackets

Surface trained	Baseline			Post-training			Training effect		
	Belly	Thigh	Shin	Belly	Thigh	Shin	Belly	Thigh	Shin
Belly	3106 (230)	3247 (320)	3286 (293)	2843 (197)	2931 (202)	2964 (199)	263 (136)	316 (221)	322 (232)
Thigh	3216 (259)	3182 (240)	3311 (266)	2966 (236)	3008 (205)	2971 (201)	250 (250)	174 (175)	340 (277)
Shin	3174 (244)	3279 (349)	3218 (336)	2863 (307)	3038 (349)	2973 (351)	311 (237)	241 (250)	245 (205)
Untrained (control)	3422 (290)	3551 (260)	3548 (293)	3225 (279)	3335 (207)	3336 (211)	197 (251)	216 (197)	212 (198)

**Table 3.**

Confusion matrices of trained participants' responses for each stimulated surface in the baseline and post-training sessions. The scores are expressed in percentages. Digits in bold indicate correct recognition percentages. Values may not add to 100% due to rounding

Stimu- lus	Baseline									Stimu- lus	Post-training								
	Response										Response								
	B	C	F	G	K	N	P	Q	Miss- ing		B	C	F	G	K	N	P	Q	Miss- ing
Belly										Belly									
B	<b>97</b>	0	0	1	0	1	0	1	1	B	<b>99</b>	0	0	0	0	0	1	0	0
C	0	<b>91</b>	1	5	0	0	2	1	1	C	0	<b>97</b>	1	1	0	0	0	1	1
F	0	0	<b>80</b>	0	11	0	8	0	1	F	1	0	<b>89</b>	1	8	0	1	0	0
G	0	2	0	<b>90</b>	2	0	2	2	2	G	0	1	1	<b>91</b>	1	1	1	6	0
K	0	0	13	1	<b>75</b>	8	1	2	1	K	1	0	4	1	<b>89</b>	5	0	1	0
N	0	0	2	3	3	<b>82</b>	7	0	3	N	0	0	0	1	1	<b>94</b>	1	1	3
P	2	0	1	0	3	2	<b>91</b>	1	2	P	0	0	1	1	1	1	<b>97</b>	1	0
Q	1	0	1	18	1	0	1	<b>79</b>	0	Q	0	0	0	4	1	0	0	<b>93</b>	1
Thigh										Thigh									
B	<b>92</b>	0	0	1	4	1	0	2	0	B	<b>99</b>	0	0	1	0	0	0	0	0
C	0	<b>89</b>	1	5	1	0	1	1	3	C	1	<b>92</b>	0	7	0	0	1	0	0
F	2	0	<b>67</b>	1	19	4	6	2	1	F	0	0	<b>80</b>	0	13	2	3	1	0
G	1	6	1	<b>79</b>	2	0	2	9	1	G	1	3	0	<b>92</b>	1	0	1	3	0
K	1	0	16	4	<b>58</b>	14	3	2	2	K	0	0	6	0	<b>86</b>	7	2	0	0
N	0	1	1	1	7	<b>79</b>	9	0	3	N	0	0	1	1	1	<b>90</b>	7	0	0
P	3	1	2	3	2	3	<b>82</b>	2	2	P	1	0	1	0	2	4	<b>89</b>	2	1
Q	3	2	1	18	1	1	1	<b>75</b>	1	Q	0	0	1	10	1	0	0	<b>88</b>	0
Shin										Shin									
B	<b>86</b>	0	0	1	1	5	2	3	3	B	<b>96</b>	0	0	1	0	1	0	1	1
C	0	<b>92</b>	0	4	0	0	1	1	1	C	0	<b>95</b>	1	1	1	1	2	0	0
F	1	0	<b>77</b>	2	15	1	2	0	2	F	1	0	<b>86</b>	0	11	1	2	0	0
G	1	10	0	<b>79</b>	1	1	2	6	2	G	1	2	1	<b>92</b>	2	0	1	1	1
K	2	0	28	2	<b>57</b>	8	1	1	2	K	0	0	13	0	<b>80</b>	6	0	1	0
N	1	1	2	3	5	<b>61</b>	21	1	6	N	0	0	2	1	1	<b>76</b>	19	1	1
P	3	0	3	1	4	4	<b>78</b>	1	6	P	2	0	1	1	1	3	<b>92</b>	0	1
Q	1	0	1	23	3	0	0	<b>72</b>	1	Q	0	0	1	8	2	1	0	<b>89</b>	0

intensive training with trial-by-trial feedback (trained group) improved performance to a greater extent than the participants merely repeating the task twice on each body surface (control group). Second, performance improvement was not restricted to the trained body surface but it transferred to the untrained ones. Third, the obtained transfer of tactile learning was similar for surfaces represented in adjacent (e.g., belly and thigh) and in non-adjacent

(e.g., belly and shin) areas of the somatosensory cortex (see Merzenich *et al.*, 1978; Nakamura *et al.*, 1998, for descriptions of the somatosensory cortex's topographic organization). Taken together, these results reveal that there is a transfer of learning in tactile letter recognition, which occurs independently of the distance between body surfaces. These three results provide support for Bach-y-Rita's claim that, after training with visual-to-tactile sensory substitution devices, perceptual abilities reached on a given body surface transfer to another one (Bach-y-Rita and Kercel, 2003). However, as will be further highlighted below, the extent to which there is no loss in perceptual performance needs to be further qualified.

The degree of learning transfer is considered to reflect the level of involved perceptual processing (Recanzone *et al.*, 1992). Transfer has therefore been reported to depend both on the task and on the stimuli. For example, transfer was not found when using a passive two-point discrimination task (Dinse *et al.*, 2006; Godde *et al.*, 2000) whereas transfer was found in gratings tasks, which involve more active discrimination, i.e. actively scanning the shapes with the fingertip (Harrar *et al.*, 2014; Sathian and Zangladze, 1997). Transfer also depends on the type of tactile stimuli (vibration, pressure, or roughness; see Harris *et al.*, 2001). Moreover, using low-level stimuli only, the above-mentioned studies reported a transfer limited to adjacent and homologous contralateral body surfaces only. The results reported here reveal that, contrary to the discrimination of low-level stimuli, the recognition of high-level symbols transfers across body surfaces without restrictions of their closeness in location. As compared with discrimination tasks of low-level stimuli varying along one feature (e.g., gratings with different orientations), the recognition of letters, which are made of several features in combination, may require the involvement of a holistic process allowing the integration of the different features into a global and unique percept. This holistic process involves higher-level areas than those perceptual processes that operate on simple features (Lerner *et al.*, 2001; Tanaka, 2003). Moreover, in our study, the sequential mode of drawing the letters may have involved a spatiotemporal integration process rather than a purely spatial one that would be involved with a simultaneous mode of presentation. This spatiotemporal integration process also depends on high-level areas (Battelli *et al.*, 2007). High-level perceptual areas are less topographically organized than low-level ones. Consequently, their receptive fields are larger and not restricted to a specific area, facilitating the transfer of perceptual learning across body surfaces that are further apart. The transfer of learning we found regarding high-level symbol recognition is thus compatible with the reverse hierarchy theory, which describe a perceptual learning guided by a top-down process, from high-level to low-level areas (Ahissar and Hochstein, 2004).

In the letter recognition task used in our study, training effects occurred at different levels. During the training session, the repetition of letters improved the recognition of the sequence of tactile stimuli corresponding to each letter. The trial-by-trial feedback provided in this training session also helped participants to associate the sequence of tactile stimuli with the correct letter. In addition, the performance improvement that was observed for the control group indicates that participants are learning the task before the training session. This result is probably influenced both by the use of three different blocks during the baseline session and by the repeated exposure of letters within each block. Harrar *et al.* (2014), using a similar procedure but with a grating discrimination task, did not find such a performance improvement for their control group. This difference between Harrar *et al.*'s and our study suggests that perceptual learning is faster for high-level than for low-level stimuli; which is consistent with the reverse hierarchy theory (Ahissar and Hochstein, 2004). However, the greater performance improvement that was obtained for the trained group than for the control group suggests that explicit training improved performance beyond a mere repeated exposure effect and thereby it supports the claim that a perceptual learning transfer occurred in our study.

It should be mentioned that the perception of the global configuration (e.g., the letter B which is made of one vertical stroke and two loops) was probably sufficient to correctly recognize tactile letters. However, the participants in our study reported more difficulties in correctly discriminating between two letters with similar global configurations (e.g., the letters F and K which are both made of a vertical stroke and two other strokes; see the confusion matrices in Table 3) that require finer discrimination processes (e.g., to discriminate between the horizontal strokes of the letter F and the diagonal strokes of the letter K). The discrimination of feature differences required for recognition of high-level symbols with high similarity may thus involve lower-level processes and reduce the amount of perceptual learning transfer. However, according to the reverse hierarchy theory, expert perceivers (in touch as well as in other sensory modalities), who have had substantial amounts of training, are able to base their perception on high-level processes even in difficult conditions (e.g., discrimination of stimuli with high similarity such as two exemplars of the same bird species); they then show transfer of perceptual learning (Ahissar and Hochstein, 2004). Intensive training should thus allow transfer of tactile learning even in the cases of discrimination of highly similar tactile symbols.

Although the results obtained in the present study support the transfer of tactile learning across body surfaces, they cannot fully support the independence of perceptual abilities from the stimulated body surface that has been claimed to occur after training. The most important counter-argument to this

claim is that there would remain differences in tactile recognition and discrimination abilities as a function of the location in which the stimuli are displayed on the body surface. For instance, tactile discrimination abilities have been reported to be higher on the belly than on the legs (Haggard *et al.*, 2003). Similarly, in our study, training improved letter recognition on all body surfaces but letter recognition remained less accurate on the legs than on the belly. However, even if the claim must be watered-down in order to include differences in acuity across body surfaces (i.e., as a function of the differences in resolution of their respective tactile receptive fields) further investigations would still be required to demonstrate an independence from the stimulated body surface. First of all, the participants in our study were given purely sensory learning whereas sensory substitution devices involve a sensorimotor learning. This however has been shown to reinforce learning transfer across different sensory modalities (see Levy-Tzedek *et al.*, 2012). Second, the body surfaces used in our study remain relatively closely represented in the somatosensory cortex; thus the use of more distant body surfaces is necessary to more fully assess the possibility of complete transfer. If perceptual learning can completely transfer to all body surfaces, this should include distant body surfaces such as head and legs. Third, the tactile letters we used are more complex than low-level stimuli varying along only one feature but most sensory substitution devices convey even higher-level stimuli such as faces, objects and scenes. Moreover, the sequential mode of presentation may be more appropriate for letters than for other objects (such as everyday 3-D objects) for which it would be difficult to sequentially order their different parts in an intuitive way. Further studies should then extend the actual results to a broader range of stimuli of varying complexity and to alternative modes of presentation.

Finally, it has been suggested that the body surface independence in visual-to-tactile sensory substitution is probably due to the fact that, after training, users no longer feel the stimulation on their skin, but they directly attribute the stimulation as resulting from an external object (see Bach-y-Rita and Kerchel, 2003). This externalization process would reinforce the independence of tactile learning from the stimulated surface (see Hartcher-O'Brien and Auvray, submitted, for a review on externalization). This claim would need further investigation as well to be corroborated. However, what can be said for now is that our study provides the first empirical data to support Bach-y-Rita's claim that training in sensory substitution improves perceptual abilities, and that these perceptual abilities transfer from one body surface to another, resulting in a relative independence from the stimulated body surface. These results have implications for the use of sensory substitution devices and vibrotactile systems by visually impaired people. First, they suggest that training on one body surface is beneficial to the entire body. As a consequence, users do not need to undergo an extensive training on all body surfaces. This study



also provides better understanding of the learning process, which is a crucial feature for efficient use of tactile devices. It reveals that transfer of learning for tactile letters occurs within a relatively short time frame, i.e., the 90 min of our experiment. The longer learning reported for recognition of more complex objects with sensory substitution devices (Kaczmarek and Haase, 2003) indicates that the amount of training required to reach accurate object recognition with tactile devices depends on the complexity of the information that is provided. It would be interesting in future research to investigate the extent to which learning transfers between different subsets of stimuli with similar level of complexity, and between stimuli of varying complexity; for instance between moderately complex stimuli such as tactile letters and highly complex stimuli such as everyday 3-D objects.

### *Acknowledgements*

We wish to thank Jessica Hartcher O'Brien for her careful reading and helpful comments on previous versions of this article. We thank Gaëtan Parseihan and Mathieu Dubois for their help in designing the apparatus. We also thank the editor, Charles Spence, and the two anonymous reviewers for their stimulating comments. G.A. is funded by a grant from the Agence Nationale de la Recherche (ANR-11-JSH2-003-1). The Institut Jean Nicod laboratory finds support in ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL.

### **Note**

- 1 Note that Bach-y-Rita also concurrently asserted that the camera could then be moved, for instance from a hand-held to a head-mounted display, without any loss in performance. However, the present study only focused on the effect of training on a change in the stimulated body surface and not on a change in the camera's position.

### **References**

- Ahissar, M. and Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning, *Trends Cogn. Sci.* **8**, 457–464.
- Arno, P., Vanlierde, A., Streef, E., Wanet Defalque, M. C., Sanabria-Bohorquez, S. and Veraart, C. (2001). Auditory substitution of vision: pattern recognition by the blind, *Appl. Cogn. Psychol.* **15**, 509–519.
- Auvray, M., Hanne-ton, S. and O'Regan, J. K. (2007). Learning to perceive with a visuo-auditory substitution system: localization and object recognition with The Voice, *Perception* **36**, 416–430.
- Auvray, M. and Myin, E. (2009). Perception with compensatory devices: from sensory substitution to sensorimotor extension, *Cogn. Sci.* **33**, 1036–1058.

- Bach-y-Rita, P. (1995). *Nonsynaptic Diffusion Neurotransmission and Late Brain Reorganization*. Demos, New York, NY, USA.
- Bach-y-Rita, P., Collins, C. C., Saunders, F., White, B. and Scadden, L. (1969). Vision substitution by tactile image projection, *Nature* **221**, 963–964.
- Bach-y-Rita, P. and Kercel, S. W. (2003). Sensory substitution and the human–machine interface, *Trends Cogn. Sci.* **17**, 541–546.
- Battelli, L., Pascual-Leone, A. and Cavanagh, P. (2007). The “when” pathway of the right parietal lobe, *Trends Cogn. Sci.* **11**, 204–210.
- Bubic, A., Striem-Amit, E. and Amedi, A. (2010). Large-scale brain plasticity following blindness and the use of sensory substitution devices, in: *Multisensory Object Perception in the Primate Brain*, J. Kaiser and M. J. Naumer (Eds), pp. 351–380. Springer, New York, NY, USA.
- Capelle, C., Trullemans, C., Arno, P. and Veraart, C. (1998). A real-time experimental prototype for enhancement of vision rehabilitation using auditory substitution, *IEEE Trans. Biomed. Eng.* **45**, 1279–1293.
- Deroy, O. and Auvray, M. (2012). Reading the world through sensory substitution devices, *Front. Theor. Philos. Psychol.* **3**, 457.
- Dinse, H. R., Kleibel, N., Kalisch, T., Ragert, P., Wilimzig, C. and Tegenthoff, M. (2006). Tactile coactivation resets age-related decline of human tactile discrimination, *Ann. Neurol.* **60**, 88–94.
- Fahle, M. E. and Poggio, T. E. (2002). *Perceptual Learning*. MIT Press, Cambridge, MA, USA.
- Gibson, E. J. (1969). *Principles of Perceptual Learning and Development*. Appleton Century Crofts, New York, NY, USA.
- Godde, B., Stauffenberg, B., Spengler, F. and Dinse, H. R. (2000). Tactile coactivation-induced changes in spatial discrimination performance, *J. Neurosci.* **20**, 1597–1604.
- Haggard, P., Taylor-Clarke, M. and Kennett, S. (2003). Tactile perception, cortical representation and the bodily self, *Curr. Biol.* **13**, 170–173.
- Haigh, A., Brown, D. J., Meijer, P. and Proulx, M. J. (2013). How well do you see what you hear? The acuity of visual-to-auditory sensory substitution, *Front. Psychol.* **4**, 330.
- Hanneton, S., Auvray, M. and Durette, B. (2010). The Vibe: a versatile vision-to-audition sensory substitution device, *Appl. Bionics Biomech.* **7**, 269–276.
- Harrar, V., Spence, C. and Makin, T. R. (2014). Topographic generalization of tactile perceptual learning, *J. Exp. Psychol. Hum. Percept. Perform.* **40**, 15–23.
- Harris, J. A., Harris, I. M. and Diamond, M. E. (2001). The topography of tactile learning in humans, *J. Neurosci.* **21**, 1056–1061.
- Hartcher O’Brien, J. and Auvray, M. (submitted). The process of distal attribution illuminated through studies of sensory substitution, *Multisens. Res.*
- Herzog, M. H. and Fahle, M. (1998). Modeling perceptual learning: difficulties and how they can be overcome, *Biol. Cybern.* **78**, 107–117.
- Iwamura, Y., Iriki, A. and Tanaka, M. (1994). Bilateral hand representation in the postcentral somatosensory cortex, *Nature* **369**, 554–556.
- Jansson, G. (1983). Tactile guidance of movement, *Int. J. Neurosci.* **19**, 37–46.
- Kaas, A. L., van de Ven, V., Reithler, J. and Goebel, R. (2013). Tactile perceptual learning: learning curves and transfer to the contralateral finger, *Exp. Brain Res.* **224**, 477–488.

- Kaczmarek, K. A. and Haase, S. J. (2003). Pattern identification and perceived stimulus quality as a function of stimulation current on a fingertip-scanned electrotactile display, *IEEE Trans. Neural Syst. Rehabil. Eng.* **11**, 9–16.
- Lerner, Y., Hendler, T., Ben-Bashat, D., Harel, M. and Malach, R. (2001). A hierarchical axis of object processing stages in the human visual cortex, *Cereb. Cortex* **11**, 287–297.
- Levy-Tzedek, S., Hanassy, S., Abboud, S., Maidenbaum, S. and Amedi, A. (2012). Fast, accurate reaching movements with a visual-to-auditory sensory substitution device, *Restor. Neurol. Neurosci.* **30**, 313–323.
- Loomis, J. M. (1974). Tactile letter recognition under different modes of stimulus presentation, *Percept. Psychophys.* **16**, 401–408.
- Meijer, P. B. L. (1992). An experimental system for auditory image representations, *IEEE Trans. Biomed. Eng.* **39**, 112–121.
- Merzenich, M. M., Kaas, J. H., Sur, M. and Lin, C.-S. (1978). Double representation of the body surface within cytoarchitectonic area 3b and 1 in “SI” in the owl monkey (*Aotus trivirgatus*), *J. Comp. Neurol.* **181**, 41–74.
- Nagarajan, S. S., Blake, D. T., Wright, B. A., Byl, N. and Merzenich, M. M. (1998). Practice-related improvements in somatosensory interval discrimination are temporally specific but generalize across skin location, hemisphere, and modality, *J. Neurosci.* **18**, 1559–1570.
- Nakamura, A., Yamada, T., Goto, A., Kato, T., Ito, K., Abe, Y., Kachib, T. and Kakigic, R. (1998). Somatosensory homunculus as drawn by MEG, *NeuroImage* **74**, 377–386.
- Parsons, L. M. and Shimojo, S. (1987). Perceived spatial organization of cutaneous patterns on surfaces of the human body in various positions, *J. Exp. Psychol. Hum. Percept. Perform.* **13**, 488–504.
- Pollok, B., Schnitzler, I., Mierdorf, T., Stoerig, P. and Schnitzler, A. (2005). Image-to-sound conversion: experience-induced plasticity in auditory cortex of blindfolded adults, *Exp. Brain Res.* **167**, 287–291.
- Recanzone, G. H., Merzenich, M. M. and Schreiner, C. E. (1992). Changes in the distributed temporal response properties of SI cortical neurons reflect improvements in performance on a temporally based tactile discrimination task, *J. Neurophysiol.* **67**, 1071–1091.
- Renier, L., Collignon, O., Poirier, C., Tranduy, D., Vanlierde, A., Bol, A., Veraart, C. and De Volder, A. G. (2005). Cross-modal activation of visual cortex during depth perception using auditory substitution of vision, *NeuroImage* **26**, 573–580.
- Sampaio, E., Maris, S. and Bach-y-Rita, P. (2001). Brain plasticity: ‘Visual’ acuity of blind persons via the tongue, *Brain Res.* **908**, 204–207.
- Sathian, K. and Zangaladze, A. (1997). Tactile learning is task specific but transfers between fingers, *Percept. Psychophys.* **59**, 119–128.
- Sekiyama, K. (1991). Importance of head axes in perception of cutaneous patterns drawn on vertical body surfaces, *Percept. Psychophys.* **49**, 481–492.
- Shaw, R. G. and Mitchell-Olds, T. (1993). ANOVA for unbalanced data: an overview, *Ecology* **74**, 1638–1645.
- Spengler, F., Roberts, T. P. L., Poeppl, D., Byl, N., Wang, X., Rowley, H. A. and Merzenich, M. M. (1997). Learning transfer and neuronal plasticity in humans trained in tactile discrimination, *Neurosci. Lett.* **232**, 151–154.

- Tanaka, K. (2003). Columns for complex visual object features in the inferotemporal cortex: clustering of cells with similar but slightly different stimulus selectivities, *Cereb. Cortex* **13**, 90–99.
- Yanagida, Y., Kakita, M., Lindeman, R. W., Kume, Y. and Tetsutani, N. (2004). Vibrotactile letter reading using a low-resolution tactor array, in: *Proc. 12th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 400–406. IEEE Computer Society.



ELSEVIER

Contents lists available at ScienceDirect

Vision Research

journal homepage: [www.elsevier.com/locate/visres](http://www.elsevier.com/locate/visres)

# Tactile recognition of visual stimuli: Specificity versus generalization of perceptual learning

Gabriel Arnold\*, Malika Auvray

Institut des Systèmes Intelligents et de Robotique, CNRS UMR 7222, Université Pierre et Marie Curie, Paris, France

## ARTICLE INFO

### Keywords:

Perceptual learning  
Generalization  
Learning strategies  
Multisensory perception  
Sensory substitution

## ABSTRACT

Sensory substitution devices aim at assisting a deficient sensory modality by means of another sensory modality. For instance, to perceive with visual-to-tactile devices, users learn to recognize visual stimuli through their tactile conversion. A crucial characteristic of learning lies in the ability to generalize, that is, the ability to extend the acquired perceptual abilities to both new stimuli and new perceptual conditions. The study reported here investigated the perceptual learning of tactile alphanumeric stimuli. The learning protocol consisted in alternating a repeated list of symbols with lists of new symbols. A first experiment revealed that, when each list consisted of 4 stimuli, recognition performance improved over time only for the repeated list. This result suggests that learning a small set of stimuli involves stimulus-specific learning strategies, preventing generalization. A second experiment revealed that increasing to six the set of learned stimuli results in higher generalization abilities. This result can be explained by greater difficulties in using stimulus-specific strategies in this case, thereby favouring the use of generalization strategies. Feature variability also appeared to be important to achieve generalization. Thus, as in visual perceptual learning, the involvement of stimulus-specific versus general strategies depends on task difficulty and feature variability. A third experiment highlighted that tactile perceptual learning generalizes to changes in orientation. These results are discussed in terms of brain plasticity as they influence the design of learning methods for using sensory substitution devices, with the aim to compensate visual impairments.

## 1. Introduction

Perceptual learning, defined as the performance improvement in perceptual tasks resulting from training (Fahle & Poggio, 2002), reveals the plasticity of humans' sensory systems. Sensory substitution is an illustrative example for which perceptual learning and plasticity play a crucial role. Sensory substitution devices aim at assisting a deficient sensory modality (e.g., vision) by means of another sensory modality (e.g., audition or touch; see Auvray & Myin, 2009, for a review). Even if the question of whether sensory substitution genuinely restores vision to their users is highly debatable, this technique provides a new way to receive sensory information from the environment. Brain plasticity allows users to have access to information in a novel way. For instance, after training with a visual-to-tactile conversion device, the tactile stimulation, initially felt on the surface of the skin, is attributed to external objects (Hartcher-O'Brien & Auvray, 2014).

Numerous studies have reported that sensory substitution devices enable blind and blindfolded sighted people to localize (e.g., Levy-Tzedek, Hanassy, Abboud, Maidenbaum, & Amedi, 2012; Renier et al.,

2005) and recognize simple shapes or complex objects (e.g., Arno et al., 2001; Auvray, Hanneton, & O'Regan, 2007; Pollok, Schnitzler, Mierdorf, Stoerig, & Schnitzler, 2005; Sampaio, Maris, & Bach-y-Rita, 2001). Most users are able, in one or two hours, to approximate objects' location and shape. However, to gain precision in these tasks and to perform more complex ones, a more intensive training is required. The length of the training is estimated at around eight hours with visual-to-tactile devices (Kaczmarek & Haase, 2003) and 10–15 with visual-to-auditory devices (Auvray, Hanneton, & O'Regan, 2007). It would be even more important to achieve a high level of expertise, for instance to be able to recognize facial expressions (Striem-Amit, Guendelman, & Amedi, 2012).

A crucial characteristic of learning lies in the ability to generalize, that is, to extend the acquired perceptual abilities to both new stimuli and new perceptual conditions. For instance, in the case of learning a sensory substitution device, generalization enables their users, trained in specific conditions and with specific shapes and objects, to extract the rules underlying perception with the device and to apply them to new objects or to previously learned objects perceived in new

\* Corresponding author.

E-mail address: [arnold@isir.upmc.fr](mailto:arnold@isir.upmc.fr) (G. Arnold).

<https://doi.org/10.1016/j.visres.2017.11.007>

Received 31 January 2017; Received in revised form 30 October 2017; Accepted 16 November 2017  
0042-6989/ © 2017 Elsevier Ltd. All rights reserved.

conditions. In other words, users do not simply associate visual stimuli to their auditory or tactile conversion, but they become able to extract the perceptual characteristics of a scene and of their constituting objects and then they are able to transfer this learning to new scenes and objects.

Studies have reported contrasting results regarding specificity versus generalization of perceptual learning (see [Fahle, 2005](#), for a review). For instance, visual perceptual learning can be highly specific to retinal position, object orientation, and object shape (e.g., [Sigman & Gilbert, 2000](#); [Yi, Olson, & Chun, 2006](#)). Tactile perceptual learning has also been reported to be specific to the trained body surface, with a transfer of learning restricted to adjacent or homologous contralateral surfaces (e.g., [Harrar, Spence, & Makin, 2014](#); [Sathian & Zangaladze, 1997](#)). On the other hand, studies have reported generalization of learning. For instance, [Furmanski and Engel \(2000\)](#) have found object-specific but size-invariant perceptual learning in visual object recognition. Perceptual learning has also been reported to transfer to non-adjacent body surfaces in tactile recognition of alphanumerical symbols ([Arnold & Auvray, 2014](#)).

The reverse hierarchy theory (RHT) of perceptual learning ([Ahissar & Hochstein, 2004](#)) has been proposed to account for these discrepant results. According to this theory, the degree of specificity depends on the difficulty of the perceptual task and on the level of cortical processing that is required to perform the task. For a difficult task, involving low-level cortical processes that are highly specific to the characteristics of the task (e.g., fine discrimination of visual orientations), perceptual learning is specific to the characteristics of the task (e.g., the visual orientations that have been discriminated) and it does not generalize to new characteristics (e.g., new orientations). On the other hand, for an easier task, involving higher-level processing, perceptual learning is less specific to the low-level characteristics of the task or of the stimuli. Task difficulty, either during the learning phase or during the evaluation of generalization, has indeed been reported to influence both the specificity and generalization of learning ([Jeter, Doshier, Petrov, & Lu, 2009](#); [Liu & Weinsall, 2000](#); [Wang, Zhou, & Liu, 2013](#)). Note however that, in some cases, generalization turns out to be easier with a difficult task than with an easy one ([Liu & Weinsall, 2000](#)). Other factors than task difficulty can play a role in influencing specificity versus generalization of learning. For instance, the variability of the learned stimuli contributes to generalization ([Hussain, Bennett, & Sekuler, 2012](#)). On the other hand, a long duration of the learning phase contributes to specificity ([Jeter, Doshier, Liu, & Lu, 2010](#)), possibly due to sensory adaptation ([Harris, Glikhsberg, & Sagi, 2012](#)).

Perceptual learning can also transfer from one sensory modality to another. For instance, learning to discriminate tactile intervals has been reported to transfer to the auditory sensory modality ([Nagarajan, Blake, Wright, Byl, & Merzenich, 1998](#)). Based on the RHT ([Ahissar & Hochstein, 2004](#)), [Proulx, Brown, Pasqualotto, and Meijer \(2014\)](#) made the hypothesis that the broad brain activation of multisensory processing reinforces generalization of learning, especially because multisensory processing involves high-level multimodal brain areas. Sensory substitution, by converting the information from one sensory modality into another, is thus an ideal candidate for observing general perceptual learning. In this field of research, [Kim and Zatorre \(2008\)](#) found that learning to recognize simple shapes and complex objects with a visual-to-auditory conversion device generalizes to new objects. This result was then reproduced with tactile-to-auditory conversion of shapes ([Kim & Zatorre, 2010](#)). Learning complex stimuli, corresponding to the auditory conversion of visual shapes, has also been reported to involve faster perceptual learning and greater amplitude of generalization to new stimuli than what occurred when learning simple stimuli ([Brown & Proulx, 2013](#)). Similarly [Auvray et al., 2007](#); see also [Auvray, 2004](#), for more details) found that learning every-day life objects transfers to some extent to the recognition of different objects belonging to the same category. Generalization to new stimuli with even more complex objects such as facial expressions has also been reported to occur

([Striem-Amit et al., 2012](#)). Finally, sensory-motor learning with a visual-to-auditory device has been reported to transfer across sensory modalities ([Levy-Tzedek, Novick et al., 2012](#)). However, if few studies have investigated the generalization of learning with visual-to-auditory devices, no study to date has investigated the generalization of learning during the use of a visual-to-tactile device.

The study reported here aims at evaluating whether learning to recognize tactile conversion of visual stimuli generalizes to new stimuli and to new perceptual conditions. The methodology is based on a recognition task of alphanumerical symbols presented on the participants' stomach by means of sequences of vibrotactile stimulations. Alphanumerical symbols are usually perceived visually, and rarely by touch. Thus the tactile recognition task requires training, similarly to what occurs when learning to use a visual-to-tactile substitution device. A similar task was previously used to show that learning to recognize tactile symbols transfers from trained to untrained body surfaces ([Arnold & Auvray, 2014](#)). In the present study, the tactile symbols were presented on the stomach as, for visual-to-tactile sensory substitution devices, this body part presents a good compromise between the size of the skin surface and its spatial resolution (see [Haggard, Taylor-Clarke, & Kennett, 2003](#), for spatial resolution of the different body surfaces). In addition, it is central on the body and, contrary to the hand, its position and orientation relative to the rest of the body does not change much. Finally, contrary to body surfaces with greater spatial resolution, in particular the fingertips and the tongue, the stomach is rarely used for everyday tactile perception (e.g., haptic object recognition) and it can thus be used for receiving additional tactile information without disturbing people's usual perception and action.

In the first two experiments reported here, the role of stimulus variability in the ability to generalize to new symbols during learning was evaluated by comparing the learning of sets of four symbols (Experiment 1; low variability) to the learning of sets of six symbols (Experiment 2; high variability). If stimulus variability influences the stimulus-specific versus generalized learning strategies, generalization should be better with high than with low stimulus variability. In the third experiment, the ability to generalize to new orientations was evaluated. Orientation is an important source of perceptual variability and changes in orientation frequently disturb the recognition of objects that are learned visually (e.g., [Arnold & Siéhoff, 2012](#)) or by touch (e.g., [Newell, Ernst, Tjan, & Bühlhoff, 2001](#)).

## 2. Experiment 1

Experiment 1 was designed to investigate whether learning to recognize tactile symbols generalizes to new symbols when learning a limited number of symbols. In order to investigate generalization to new symbols, a similar procedure to the one used by [Dutilh, Krypotos, and Wagenmakers \(2011\)](#) was used. It consists in alternating blocks with a repeated list of stimuli and blocks with new lists of stimuli. If stimulus-specific learning is involved, recognition performance should increase with practice (i.e., across blocks) only for repeated stimuli. In contrast, if generalized learning is involved, recognition performance should increase for both repeated and new stimuli.

The second aim of this experiment was to investigate whether the generalization of learning depends on the spatial perspective that is naturally adopted when interpreting tactile information presented on the body. Using the graphesthesia task, which consists in recognizing ambiguous asymmetrical symbols presented on the body surface (e.g., the letters b, d, p, and q), previous studies have shown that some observers prefer to adopt a self-centred spatial perspective (i.e., centred on the observer's body) whereas others prefer to adopt a decentred one (i.e., centred on a location different from that of the observer) (see [Arnold, Spence, & Auvray, in press](#), for a review). These individual differences have been reported to reflect the existence of a natural perspective rather than being due to an arbitrary choice ([Arnold, Spence, & Auvray, 2016](#)). Moreover, decentred observers are better at

adopting an unnatural perspective than self-centred observers are. Thus, this experiment also investigated whether this superiority of naturally decentred observers over naturally self-centred ones is specific to the adoption of spatial perspectives or whether it reflects a general ease to interpret tactile information instead. When tactile stimuli are presented on people's stomach, three perspectives can be adopted, one decentred and two self-centred (trunk-centred or head-centred). In the present experiment, only those participants that naturally adopt a trunk-centred or a decentred perspective were tested.

## 2.1. Material and methods

### 2.1.1. Participants

Two hundred and fifty-five participants (154 females and 101 males; mean age = 25.2 years, range = 18–47 years; 30 participants were left-handed and 225 participants were right-handed) were initially recruited via mailing lists to participate in 6 different experiments (i.e., the 3 experiments reported in the present study and 3 experiments investigating different research questions reported in different studies). In all these experiments, the graphesthesia task was used to investigate the participants' spatial perspectives. The participants were randomly allocated to one of these 6 experiments in order to generate groups based upon natural perspective. Five participants were excluded from the set of experiments because their natural perspective could not be correctly identified. For the 250 remaining participants, 49.6% adopted a trunk-centred perspective, 29.2% a head-centred perspective, and 21.2% a decentred perspective. For the present study, 20 trunk-centred and 20 decentred participants were included in Experiment 1, 32 trunk-centred participants in Experiment 2, and 24 trunk-centred participants in Experiment 3.

Forty participants completed the present experiment (25 females and 15 males; mean age = 24.8 years, range = 18–41 years; three participants were left-handed and 37 participants were right-handed). Half of the participants (14 females and 6 males; mean age = 24.4 years, range = 18–41 years) had a natural trunk-centred perspective and the other half (11 females and 9 males; mean age = 25.2 years, range = 20–33 years) had a natural decentred perspective. All the participants were naive to the purpose of the experiment. Participants provided their informed consent and received payment for their participation. The experiment took approximately one hour to complete and was performed in accordance with the ethical standards laid down in the [Declaration of Helsinki \(1991\)](#).

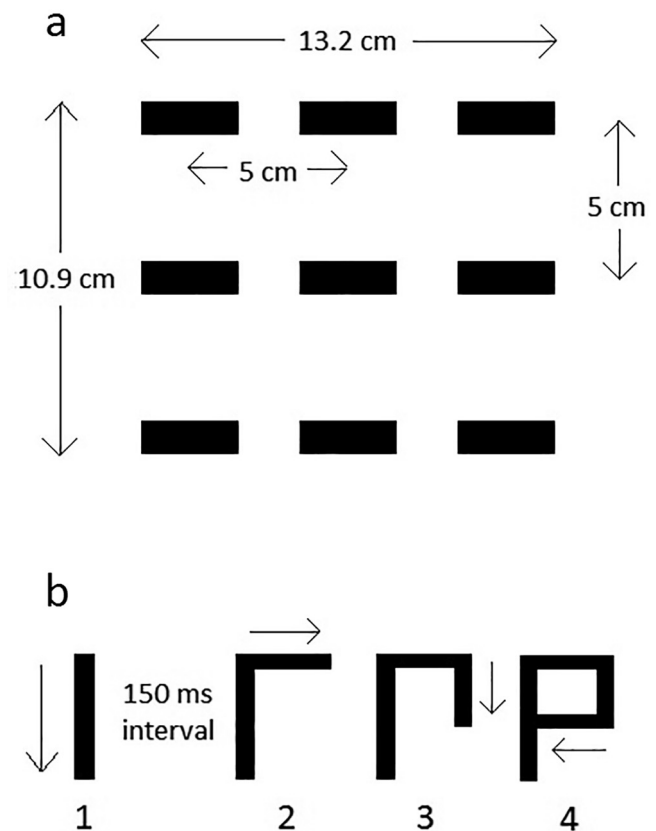
### 2.1.2. Apparatus

The tactile stimuli were presented by means of 9 rectangular vibrators (Haptuator Mark II, Tactile Labs, Montreal, Canada) arranged in a 3-by-3 array with a centre-to-centre spacing of 5 cm (see [Fig. 1a](#)). Each vibrator measured 0.9 cm vertically by 3.2 cm horizontally for a total surface area of 10.9 cm vertically by 13.2 cm horizontally. A nine-channel amplifier drove each vibrator independently at a 250-Hz frequency. The vibrator array was placed on the participants' stomach, above their clothes, by means of an elastic belt. The vibrator array was placed symmetrically to their body mid-sagittal line, with the 3 lower vibrators located above the waist. Only one layer of clothing was allowed between the skin and the vibrators. The participants individually selected the intensity of each vibrator by means of an adjustment method.

Instructions and feedback were presented on a 23-inch screen with a 1920 × 1080 resolution. The participants wore noise-reducing headphones with a noise reduction rating of 30 dB, in order to mask any sounds made by the vibrators.

### 2.1.3. Stimuli

Twenty asymmetrical symbols were presented on the participants' stomach: the uppercase letters B, C, D, E, F, G, J, K, L, N, P, Q, R, and Z and the numbers 1, 2, 3, 4, 5, and 6. The use of asymmetrical symbols



**Fig. 1.** Illustration of the drawing of alphanumerical symbols with the 3-by-3 vibrotactile matrix. (a) Schematic figure illustrating the vibrotactile matrix. (b) The sequence of vibrations chosen for drawing the letter P.

allowed presenting the symbols according to either trunk-centred or decentred perspectives. According to the trunk-centred perspective, the left and right of the symbols were congruent to participants' left-right axis. According to the decentred perspective, the left and right were opposite to participants' left-right axis. The top and bottom of the symbols were congruent to participants' top-bottom axis for both the trunk-centred and decentred perspectives.

The symbols were traced with sequences of vibrotactile stimulations mapping the trajectory of manual drawing (see [Arnold & Auvray, 2014](#), for more details). Each vibrator composing the sequence was activated for 250 ms and there were 150 ms intervals between two consecutive vibrations that were spatiotemporally discontinuous (e.g., the end of the vertical descending stroke and the beginning of the loop of the letter P; see [Fig. 1b](#)). This interval avoids possible errors of interpretation caused by an irrelevant grouping of vibrations. The participants were instructed that these temporal intervals corresponded to a spatial discontinuity in manual letter drawing. It should be underlined that with such a device, the symbols' drawings do not exactly match manual drawing, as curved features are not possible and thus they are replaced by right-angled features. Tactile recognition of alphanumerical symbols has nonetheless previously been reported to be successful with the use of a similar 3-by-3 vibrotactile matrix ([Yanagida, Kakita, Lindeman, Kume, & Tetsutani, 2004](#)). The mean duration of symbols was 2215 ms (SD = 618). Five lists of four symbols were created with similar mean durations (List1: PQZ5, 2363 ms; List2: JK36, 2175 ms; List3: FLR4, 2063 ms; List4: BCD2, 2263 ms; EGN1, 2213 ms). Each of these lists was composed of three letters and one number except List2, which was composed of two letters, and two numbers.

In addition, the four ambiguous lowercase letters b, d, p, and q were used for the identification of each participant's natural perspective. The letters were traced beginning from the stem and ending with the loop,

with a sequence of eight 250 ms vibrations without intervals between each vibration.

#### 2.1.4. Procedure

The first part of the experiment was designed to identify each participant's natural perspective. The participants sat in front of the computer screen. On each trial, one of the four possible ambiguous letters (b, d, p, and q) was displayed on the participants' stomach. The participants were instructed to report the letter they perceived as spontaneously as possible by pressing the corresponding key with the index finger of their dominant hand. They were informed that each vibrotactile sequence could be interpreted as one of the four letters, depending on how they assign the left–right and top–bottom axes of the letter, and that there were consequently no correct or incorrect responses. The participants were asked to report their responses as fast as possible. They completed one block of 16 trials (four trials for each of the four letters). If no clear preference for adopting one of the perspectives was identified in the first block, the participants completed additional blocks until their natural perspective could be clearly identified, without exceeding six blocks. Only one block was required for 24 participants, two blocks for 12 participants, and three blocks for four participants. There are three possible causes for this individual variability that can occur individually or in combination. First, the natural perspective may emerge faster for some participants over others. Second, some people might not have a natural perspective. Third, some participants may have had difficulties to understand the instructions, leading them to adopt the most 'logical' perspective (i.e., the experimenter's perspective) rather than the easiest perspective for them.

The second part of the experiment was designed to investigate participants' abilities to recognize unambiguous asymmetrical symbols. On each trial, one symbol was presented on the participants' stomach. The participants were instructed to report which symbol was presented by pressing the corresponding key with the index finger of their dominant hand. Rather than using the real position of each letter and number on the computer keyboard, stickers with the four symbols used in each list written on it were placed on four adjacent keys on the keyboard. The keyboard was located in front of the participants, slightly shifted toward the left for left-handed participants or to the right for right-handed participants. In this part of the experiment, the participants were informed that the symbols would be presented according to their natural perspective and that there were correct and incorrect responses. The participants were asked to give their responses as accurately as possible. Response times (RTs) were recorded as a secondary measure of performance but emphasis was put on accuracy (see [Supplementary material](#) for the analyses of RTs). The participants were able to give their responses at any time from the onset of the first vibration and up to 3000 ms after the end of the last vibration (at which point the trial was terminated). At the end of each trial, there was an interval of 3000 ms before the beginning of the next trial.

The participants completed eight blocks of trials for the recognition task. In each block, one of the five lists was presented. One of the lists (the repeated list; labelled A) was presented four times. The other four lists (new lists; labelled B through E) were presented once each. Thus there were four blocks with the repeated list and four blocks with the new lists, which were alternated as follows: ABACADAE for half of the participants and BACADA EA for the other half. Note that each of the five lists served as the repeated list for a fifth of the participants ( $N = 16$ ). In addition, each list was presented the same number of times across participants as lists B, C, D, and E.

Each block consisted of 40 trials (10 presentations of each symbol). At the end of each block of trials, global feedback indicating the percentage of correct responses and the mean RT was given. The participants had a short pause between each block. At the beginning of each block, the participants were given four practice trials (one trial for each symbol). In this practice block, feedback was given at the end of each trial: the correct answer was visually displayed on the computer screen

once the participants responded.

#### 2.2. Results

Trials in which the participants failed to make a response before the trial was terminated (0.56% of trials) were not included in the data analyses. The accuracy of symbol recognition was measured by the percentage of correct responses in each condition. An ANOVA was conducted with Type of list (repeated, new) and Block (block 1, block 2, block 3, block 4) as within-participant factors and with Natural perspective (trunk-centred, decentred) as a between-participant factor. As there was no significant effect of Gender and no interaction between Gender and any of the other factors (in this experiment as well as in Experiments 2 and 3), this factor was not included in the reported analyses.

There was a significant effect of Block,  $F(3,114) = 3.46$ ,  $p < .05$ , and a significant interaction between Block and Type of list,  $F(3,114) = 5.40$ ,  $p < .01$ , showing an increase in accuracy across blocks for the repeated list (block 1 = 91.2%, SD = 9.4; block 2 = 94.1%, SD = 7.2; block 3 = 96.4, SD = 5.7; block 4 = 97.6, SD = 2.6) but not for the new lists (block 1 = 92.4%, SD = 8.8; block 2 = 91.3%, SD = 10.3; block 3 = 93.1, SD = 7.3; block 4 = 92.0, SD = 7.6) (see [Fig. 2](#)). The linear trend test was significant for the repeated list,  $F(1,38) = 23.17$ ,  $p < .001$ , but not for the new lists,  $F(1,38) < 1$ . There was also a significant effect of Type of list,  $F(1,38) = 13.59$ ,  $p < .001$ , with greater accuracy for repeated (94.8%, SD = 7.1) than for new lists (92.2%, SD = 8.6). Finally, the three-way interaction between Natural perspective, Type of list, and Block almost reached significance,  $F(3,114) = 2.56$ ,  $p = .058$ . However, the analysis of this interaction revealed no difference between trunk-centred and decentred participants in generalization of learning as the linear trend tests revealed significant increase in accuracy across blocks for the repeated list but not for the new lists for the two groups (repeated list:  $F(1,38) = 11.79$ ,  $p < .01$ , for trunk-centred,  $F(1,38) = 11.38$ ,  $p < .01$ , for decentred; new lists:  $F(1,38) < 1$ , for both trunk-centred and decentred).

The number of blocks required to identify the participants' natural perspective individually varied. This might indicate either that the adopted perspective is more or less natural or that the natural perspective emerges faster for some participants over others. This variability was taken into account in order to further investigate the influence of natural perspective on learning and generalization. More specifically, the participants who required only one block of ambiguous symbols ( $N = 24$ ) were compared with the participants who required

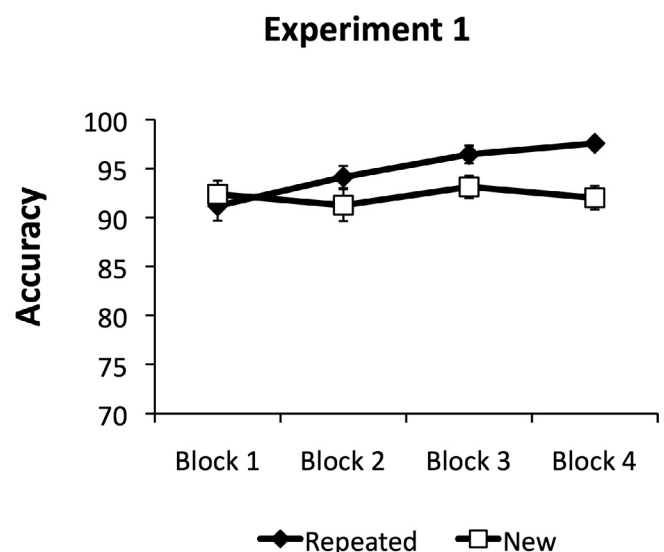


Fig. 2. Participants' accuracy during Experiment 1 as a function of Type of List (repeated, new) and Block. Error bars represent the standard errors of the means.



several blocks (i.e., two or three blocks;  $N = 16$ ). An ANOVA was conducted with Number of blocks (one, several) as a between-participant factor and with Type of list (repeated, new) and Blocks (block 1, block 2, block 3, block 4) as within-participant factors. The results revealed a significant effect of Number of blocks,  $F(1,38) = 6.73$ ,  $p < .05$ , with greater accuracy for participants with one block (95.0%,  $SD = 3.3$ ) than for participants with several blocks (91.2%,  $SD = 6.0$ ). They also revealed a significant interaction between this factor and Blocks,  $F(3,114) = 4.03$ ,  $p < .01$ , showing a greater increase in accuracy for participants with several blocks (from 87.5% in block 1 to 91.9% in block 4) than for participants with one block (from 94.6% in block 1 to 95.1% in block 4). However, there was no significant interaction between Number of blocks and Type of list,  $F(1,38) < 1$ ,  $ns$ , and no significant interaction between Number of blocks, Type of list, and Blocks,  $F(3,114) = 2.00$ ,  $p = .118$ . Finally, linear trend tests indicated a significant increase in accuracy across blocks for the repeated list but not for the new lists. This was the case both for participants with one block ( $F(1,23) = 14.07$ ,  $p < .01$  for repeated;  $F(1,23) = 2.29$ ,  $p = .144$  for new) and for participants with several blocks ( $F(1,15) = 15.060$ ,  $p < .01$  for repeated;  $F(1,15) = 2.97$ ,  $p = .106$  for new).

### 2.3. Discussion

The results of Experiment 1 showed that tactile symbol recognition improved with training. However, performance improvement was restricted to those symbols that were repeated throughout the experiment, and did not occur for the new symbols that were introduced. The participants may thus have used stimulus-specific strategies of learning rather than extracting the perceptual rules that enable generalization of learning. In this experiment, the use of only four different tactile symbols in each block may have biased the participants toward adopting stimulus-specific strategies. For instance, it is possible to identify local (e.g., the presence of a curve for only one symbol in a list) or temporal (e.g., the duration of one symbol is shorter than the others) cues to recognize each symbol. With only four symbols the cost of maintaining such specific cues in short-term memory is not too important. However, the counterpart of such stimulus-specific strategies is that the cues specific to the learned symbols are no longer relevant when new symbols are introduced. As stimulus variability during learning facilitates generalization (Hussain et al., 2012), Experiment 2 therefore investigated whether learning a greater number of stimuli in each block encourages the participants to adopt generalized learning strategies.

Experiment 1 was also designed to evaluate whether tactile symbol recognition depends on the participants' natural preference for adopting self-centred or decentred perspectives. The results did not reveal any differences in recognition performance between the participants naturally adopting trunk-centred and those adopting decentred perspectives. This was true even when the variability in the emergence of the natural perspective was taken into account. Thus, both groups have similar results when participants were asked to recognize unambiguous symbols, contrary to the results of a previous study, in which decentred participants were better at recognizing ambiguous symbols (b, d, p, and q) from an unnatural perspective than self-centred participants (Arnold et al., 2016). The superiority of decentred observers in the recognition of ambiguous symbols may consequently be specific to assigning spatial coordinates (left, right, top, and bottom) to tactile stimulation and not to shape processing *per se*. This specificity is in line with the dissociation between processing object identity and processing object orientation, which has been suggested to rely on the ventral and dorsal visual streams respectively (Valyear, Culham, Sharif, Westwood, & Goodale, 2006).

## 3. Experiment 2

In Experiment 1, tactile symbol recognition only improved across blocks for repeated symbols but not for new symbols, revealing stimulus-specific rather than generalized learning. This second experiment investigated whether the learning strategy involved depends on the quantity of information to-be-learned and whether learning a greater number of symbols in each block favours a generalized learning. There were 6 symbols to learn in each list, with the same alternation of repeated and new symbols as in Experiment 1. If a more generalized learning is involved here, a symbol recognition improvement for both repeated and new symbols should be observed. As the recognition of unambiguous symbols did not depend on the participants' natural perspective in Experiment 1, and as naturally trunk-centred participants are more frequent than others (Arnold et al., 2016), only naturally trunk-centred participants were tested in this experiment. The symbols were always presented according to the participants' natural trunk-centred perspective.

### 3.1. Material and methods

#### 3.1.1. Participants

Thirty-two participants (27 females and 5 males; mean age = 22.9 years, range = 18–40 years; eight participants were left-handed and 24 participants were right-handed) that all had a natural trunk-centred perspective completed the experiment. All the participants were naive to the purpose of the experiment. None had participated in the first experiment. Participants provided their informed consent and received payment for their participation. The experiment took approximately one hour to complete and was performed in accordance with the ethical standards laid down in the Declaration of Helsinki (1991).

#### 3.1.2. Apparatus

The apparatus was exactly the same as in Experiment 1.

#### 3.1.3. Stimuli

The 20 symbols used in Experiment 1 and four additional symbols (the uppercase letters H, S, T, and U) were presented on the participant's stomach. The way the letters were traced was the same as in Experiment 1. The mean duration of symbols was 2188 ms ( $SD = 586$ ). Four lists of six symbols were created with similar mean durations (List1: JKPZ36, 2142 ms; List2: FHLRT4, 2075 ms; List3: BCDQ25, 2392 ms; List4: EGNSU1, 2142 ms). Each of these lists was composed of four or five letters and one or two numbers.

#### 3.1.4. Procedure

The procedure was similar to the one used in Experiment 1. In the first part of the experiment, in order to decrease the number of blocks necessary to identify their natural perspective, participants were explicitly instructed to adopt the easiest perspective for them rather than the most 'logical' perspective. To identify each participant's natural perspective, only one block was required for 25 participants, two blocks for two participants, and three blocks for five participants.

Once the participants' natural perspective was identified, they completed six blocks for the recognition task. One of the lists (repeated list; labelled A) was presented three times. The other three lists (new lists; labelled B through D) were presented once each. Therefore there were three blocks with the repeated list and three blocks with the new lists, which were alternated as follows: ABACAD for half of the participants and BACADA for the other half. Each of the four lists served as the repeated list for a fourth of the participants ( $N = 8$ ). In addition, each list was presented the same number of times across participants as lists B, C, and D. Each block consisted of 60 trials (10 presentations of each symbol). At the beginning of each block, the participants were given six practice trials (one trial for each symbol) with a feedback

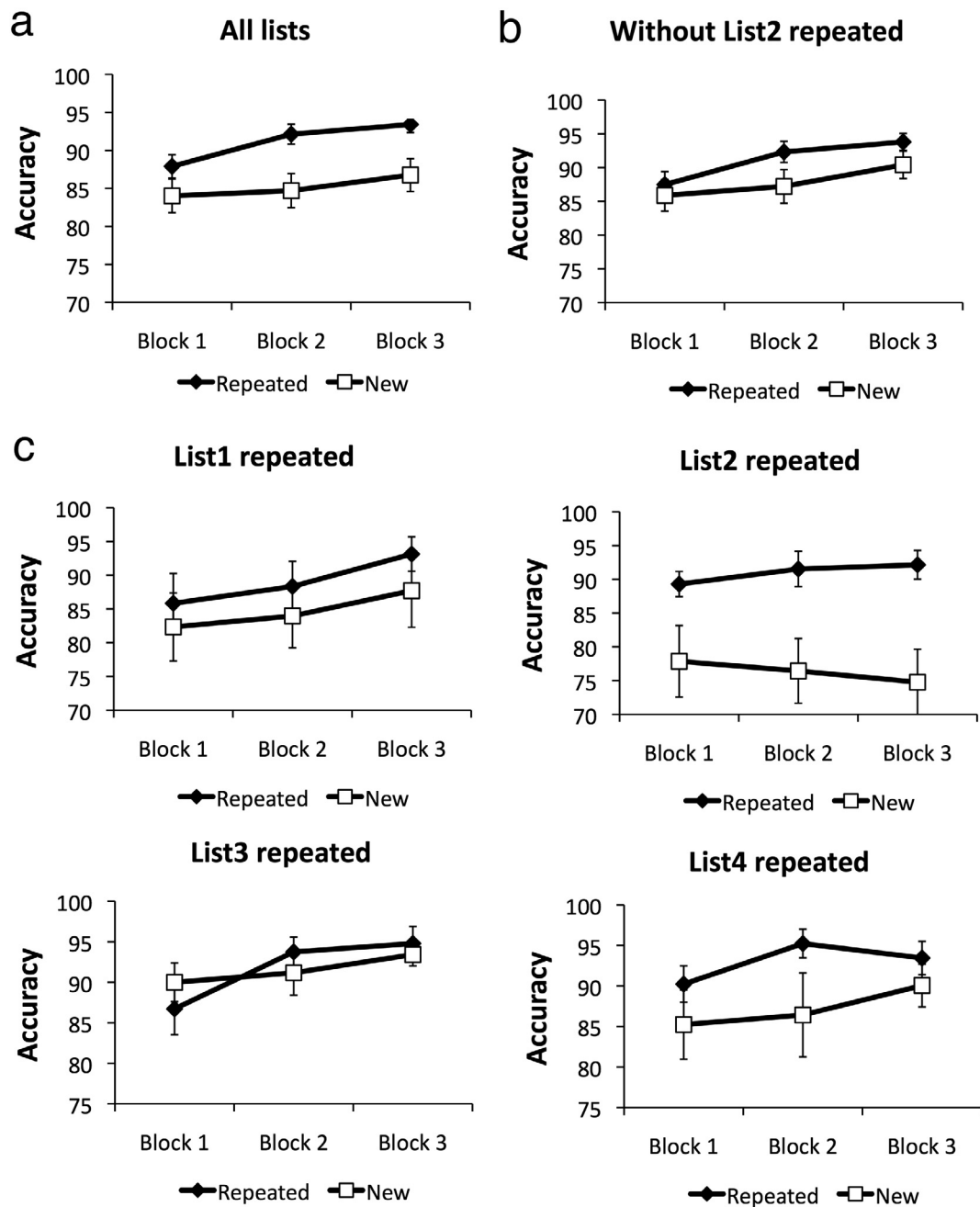


Fig. 3. Participants' accuracy during Experiment 2 as a function of Type of list (repeated, new) and Block. (a) When all the lists were included, accuracy increased across blocks for the repeated list only. (b) When excluding the participants who had List 2 as the repeated list, accuracy increased across blocks for both the repeated and new lists. (c) Details of the results for each list as the repeated list. Accuracy increased across blocks for the new lists except when List 2 was the repeated list. Error bars represent the standard errors of the means.

indicating the correct response at the end of each trial.

### 3.2. Results

Trials in which the participants failed to make a response before the trial was terminated (2.00% of trials) were not included in the data analyses. An ANOVA was conducted with Type of list (repeated, new) and Block (block 1, block 2, block 3) as within-participant factors. The results showed a significant effect of Block,  $F(2,62) = 8.27, p < .001$ , and, contrary to Experiment 1, no significant interaction between Block and Type of list,  $F(2,62) = 1.51, p = .230$ . This reveals a global increase in accuracy across blocks (block 1 = 86.5%, SD = 10.8; block 2 = 88.8%, SD = 11.0; block 3 = 90.5, SD = 10.1) (see Fig. 3a). The increase in accuracy was observed for both repeated (block 1 = 88.5%,

SD = 8.6; block 2 = 92.5%, SD = 7.5; block 3 = 93.7, SD = 6.0) and new lists (block 1 = 84.5%, SD = 12.5; block 2 = 85.0%, SD = 12.8; block 3 = 87.2, SD = 12.3). The linear trend test was however significant for the repeated list,  $F(1,31) = 18.12, p < .001$ , but not for the new lists,  $F(1,31) = 1.97, p = .171$ . There was also a significant effect of the Type of list,  $F(1,31) = 20.40, p < .001$ , with greater accuracy for repeated (91.5%, SD = 7.7) than for new lists (85.6%, SD = 12.4).

The absence of interaction between Type of List and Block suggests a global increase in accuracy across blocks but, nevertheless, there was no significant linear increase for the new lists. Thus, to try and account for the results of Experiment 2, the effect of each of the lists was investigated, to see whether the generalization of learning depends on the learned symbols. There was an increase in accuracy across blocks with the new lists when all the lists were repeated except for List 2 (see

**Table 1**  
Number of horizontal, vertical, oblique, and curved features for the different lists in Experiment 1 and Experiment 2.

		Horizontal	Vertical	Oblique	Curve	Total
Experiment 1	List 1	3	2	2	3	10
	P Q Z 5					
	List 2	2	2	2	4	10
	J K 3 6					
	List 3	4	5	1	1	11
	F L R 4					
	List 4	1	2	1	5	9
	B C D 2					
	List 5	4	4	3	1	12
	E G N 1					
Experiment 2	List 1	4	3	3	5	15
	J K P Z 3 6					
	List 2	6	8	1	1	16
	F H L R T 4					
	List 3	2	3	2	7	14
	B C D Q 2 5					
	List 4	4	4	3	4	15
	E G N S U 1					

Fig. 3c). The difference in accuracy between the repeated and the new lists was also the strongest when List 2 was repeated (13.4 points of percentage; 4.4, 0.2, and 5.8 points of percentage for the List 1, List 3, and List 4, respectively), showing strong specificity to the stimulus learned in this list. It is interesting to note that List 2 differed from the others by the variability in its letter features. Table 1 shows that the symbols composing List 2 were essentially made of vertical and horizontal features and that there were very few oblique and curved features. On the contrary, the different types of features were more equally distributed in List 1, List 3, and List 4. Thus, feature variability may also have played a role in the generalization of learning, with exposition to a small set of features (List 2) potentially limiting one's ability to recognize new symbols.

Accordingly, an ANOVA was conducted, this time without those participants who had List 2 as the repeated list. In this case, the interaction between Type of List and Block was not significant,  $F(2,62) = 1.34$ ,  $p = .271$ , and the linear trend test was significant for both the repeated,  $F(1,31) = 15.69$ ,  $p < .001$ , and the new lists,  $F(1,31) = 5.93$ ,  $p < .05$ , showing a clear pattern of generalization (see Fig. 3b). Note that this was not the case when the participants who had the List 1, List 3, or List 4 as the repeated list were excluded from the analysis. In order to evaluate the role of feature variability in Experiment 1, the distribution of the different types of features in the different lists was also analysed. Table 1 shows that the different types of features were relatively well distributed in List 1, List 2, and List 5 whereas List 3 was essentially made of vertical and horizontal features and List 4 was essentially made of curved features. Thus an ANOVA was conducted with only the participants who had List 1, List 2, and List 5 as the repeated list. The results revealed a significant interaction between Type of List and Block,  $F(3,69) = 5.98$ ,  $p < .01$ , and the linear trend test was significant only for the repeated list,  $F(1,23) = 18.96$ ,  $p < .001$ , not for the new lists,  $F(1,23) = 1.75$ ,  $p = .199$ , showing no pattern of generalization. In summary, the results of Experiment 2 indicate that participants implemented general learning strategies except when feature variability of the repeated list was relatively low. However, in Experiment 1, with a few learned symbols, stimulus-specific strategies were involved, independently of feature variability.

### 3.3. Discussion

Compared to the results of Experiment 1, the results of Experiment 2 are overall in favour of generalized learning strategies and they are compatible with the effect of stimulus variability on generalization (Hussain et al., 2012). When the number of symbols to learn is

increased (from four to six symbols in each list), stimulus-specific strategies may be too costly. For instance, it may be difficult to keep a specific spatio-temporal cue for each symbol to learn in working memory. As a consequence, participants try to extract general perceptual rules, which favour the subsequent recognition of unlearned stimuli. An alternative explanation would be that task difficulty influenced the degree of generalization. Indeed, the global recognition performance was significantly higher in Experiment 1 (mean accuracy = 93.5%,  $SD = 4.8$ ; mean RT = 2420 ms,  $SD = 360$ ), with four symbols to recognize in each block, than in Experiment 2 (mean accuracy = 88.6%,  $SD = 8.4$ ; mean RT = 2743 ms,  $SD = 287$ ), with six symbols, for both accuracy,  $t(70) = 3.14$ ,  $p < .01$  and RTs,  $t(70) = 4.10$ ,  $p < .001$ . However, contrasting results have been reported for the effect of task difficulty, with some studies showing specificity for difficult tasks (Jeter et al., 2009; Wang et al., 2013) but other showing generalization (Liu & Weinsall, 2000).

The results of Experiment 2 also revealed that generalization depends on the symbols' feature variability. For one list of symbols among the four different lists, perceptual learning remained specific to the learned symbols. Compared to the other lists, this list consisted of symbols with a low variability, essentially made of vertical and horizontal strokes, with few oblique or curved features. However, the same analysis of feature variability conducted on the results of Experiment 1 did not show any influence of feature variability: stimulus-specific learning was observed for both homogeneous and heterogeneous lists of symbols. Overall, the results of Experiment 1 and Experiment 2 therefore show that generalization to new symbols depends both on the size of the stimulus set that is learned and on feature variability. When fewer symbols are learned, stimulus-specific strategies are involved, independently of feature variability. When the stimulus set is increased, generalized strategies are involved but feature variability is necessary to make these strategies efficient. This effect of variability during training is consistent with previous studies reporting greater generalization after learning objects from different exemplars (Baecck, Maes, Van Meel, & Op de Beeck, 2016).

Finally, one possibility is that the type of learned features also plays a role. It may be more difficult to generalize from horizontal and vertical to oblique and curved features than the other way round. This effect of symbol feature can be interpreted with the feature analysis theory of visual perception (Treisman & Gormican, 1988). According to this theory, basic visual features such as straight lines are coded in the visual brain as standards and other features such as oblique and curved lines are coded as a deviation from the standard. As a consequence, brain activation is more important when perceiving deviant than standard features. A possibility here is that tactile recognition of visual symbols involved similar low-level processes as in visual recognition and that learning heterogeneous symbols, made of horizontal, vertical, oblique, and curvature features, activated broader brain areas than learning homogeneous symbols. According to the RHT, this broader brain activation has subsequently facilitated learning generalization.

### 4. Experiment 3

Experiment 3 was designed to investigate generalization to changes in orientation. Tactile symbols were learned upright and then they had to be recognized rotated 90° to the left or to the right, or 180° (upside-down). In order to be able to directly compare the generalization to new stimuli and to new orientations, the same learning protocol as in Experiments 1 and 2 was used, with an alternation of repeated (upright) and new (rotated) orientations. If learning generalizes to changes in orientation, that is, if recognizing a learned symbol in a new orientation has no cost, an improvement in recognition performance for both the repeated and new orientations should be observed. On the contrary, if learning is specific to the learned orientation, an improvement in performance only for the repeated orientation should be observed instead. As in Experiment 2, only naturally trunk-centred participants

completed this experiment.

#### 4.1. Material and methods

##### 4.1.1. Participants

Twenty-four participants (16 females and 8 males; mean age = 23.1 years, range = 19–34 years; one participant was left-handed and 23 participants were right-handed) that all had a natural perspective centred on their trunk completed the experiment. All the participants were naive to the purpose of the experiment. None had participated in Experiment 1 or Experiment 2. Participants provided their informed consent and received payment for their participation. The experiment took approximately one hour to complete and was performed in accordance with the ethical standards laid down in the [Declaration of Helsinki \(1991\)](#).

##### 4.1.2. Apparatus

The apparatus was the same as in Experiments 1 and 2.

##### 4.1.3. Stimuli

The same four lists of six symbols as presented in Experiment 2 were used. For each symbol, one vibrotactile sequence was created for each orientation (upright, left-rotated, right-rotated, upside-down). The left-rotated orientation corresponded to a 90°-rotation of the symbol toward the participant's left-hand side. Therefore the top of the symbol was located on the left side of the trunk. In contrast, the right-rotated orientation corresponded to a 90°-rotation of the symbol toward the participant's right-hand side. Therefore the top of the symbol was located on the right side of the trunk. The upside-down condition corresponded to a 180°-rotation of the symbol, resulting in the top of the symbol being located on the bottom of the trunk.

##### 4.1.4. Procedure

The procedure was similar to the one of Experiment 2. To identify each participant's natural perspective, only one block was required for 22 participants, two blocks for one participant, and three blocks for one participant.

After the identification of their natural perspective, the participants completed six blocks for the recognition task. For each participant, only one list was presented, with the same number of participants ( $N = 6$ ) for each list. In each block, the symbols were presented with one out of the four orientations. The upright orientation (repeated, labelled A) was presented three times. The other three orientations (unrepeated, labelled B through D) were presented once each. Thus there were three blocks with the symbols presented upright and three blocks with the symbols rotated, that were alternated as follows: ABACAD for half of the participants and BACADA for the other half. Each rotated condition (90° left, 90° right, 180°) was presented the same number of times across participants as blocks B, C, and D. Each block consisted of 60 trials (10 presentations of each symbol). At the beginning of each block, the participants were given six practice trials (one trial for each symbol) with feedback indicating the correct response at the end of each trial.

#### 4.2. Results

Trials in which the participants failed to make a response before the trial was terminated (0.26% of trials) were not included in the data analyses. An ANOVA was conducted with Type of orientation (repeated, new) and Block (block 1, block 2, block 3) as within-participant factors. The results showed a significant effect of Block,  $F(2,46) = 9.67$ ,  $p < .001$ , and no significant interaction between Block and Type of orientation,  $F(2,46) < 1$  (see [Fig. 4a](#)). The increase in accuracy was observed for both repeated (block 1 = 84.4%,  $SD = 13.4$ ; block 2 = 88.2%,  $SD = 13.2$ ; block 3 = 90.4%,  $SD = 10.7$ ) and new (block 1 = 77.9%,  $SD = 16.5$ ; block 2 = 81.8%,  $SD = 14.6$ ; block 3 = 82.8%,  $SD = 16.1$ ) orientations. The linear trend test was

significant for both the repeated orientation,  $F(1,23) = 12.48$ ,  $p < .01$ , and the new orientations,  $F(1323) = 5.88$ ,  $p < .05$ . There was also a significant effect of Type of orientation,  $F(1,23) = 25.65$ ,  $p < .001$ , with greater accuracy for repeated (87.7%,  $SD = 12.6$ ) than for new orientations (80.8%,  $SD = 15.7$ ).

In order to investigate a possible effect of the symbols' orientation on learning generalization, the accuracy was also computed as a function of the angle of rotation (90° left, 90° right, and 180°) from the upright orientation (see [Fig. 4b](#)). The results first revealed that global accuracy was similar for the three angles of rotation (81.3%,  $SD = 16.5$ , for the 90° left; 81.3%,  $SD = 15.0$ , for the 90° right; 80.0%,  $SD = 16.2$ , for the 180°). It should be underlined that this experiment was designed to investigate the participants' ability to recognize learned symbols with a new orientation in each session. As a consequence, the evaluation of performance improvement across blocks as a function of each symbol's orientation is based on between-participant comparisons with a small sample size ( $n = 6$ ) and can be only descriptive. [Fig. 4b](#) illustrates that recognition performance improved across blocks for all the symbols' orientations except for the 90°-left orientation for which performance increased in the second block and then decreased in the third block.

#### 4.3. Discussion

The results of Experiment 3 revealed generalization to previously learned symbols presented in a new orientation and they are compatible with previously reported results on generalization to new orientations. Even though perceptual learning is often highly specific to orientation (e.g., [Hussain, Sekuler, & Bennett, 2009](#)), generalization to new orientations has been reported to be easier than generalization to new set of stimuli ([Baeck, Windey, & Op de Beeck, 2012](#)). The sequential presentation of the tactile symbols in our study may have also facilitated generalization to orientation. Indeed, the spatio-temporal characteristics of each symbol may easily be recognized independently of the way the symbol is oriented on the body surface. For instance, the sequence of one long stroke followed by two short strokes for the letter K is easily perceived in each orientation. That is to say, the static presentation of a symbol might be seen to facilitate generalization to the same extent.

The results of Experiment 3 are also compatible with previous studies on visual recognition of rotated letters and digits, which showed that recognition depends on angular rotation as a function of the task. When the task consists in deciding whether the symbol is normal or mirror-reversed, there is an effect of angular rotation on recognition performance ([Cooper & Shepard, 1973](#)). On the contrary, when the task consists in identifying the alphanumeric symbol, there is no effect of angular rotation on recognition performance, showing that alphanumeric symbols are recognized through the extraction of critical features invariant to the symbol's orientation ([Corballis, Zbrodoff, Shetzer, & Butler, 1978](#); [Eley, 1982](#)). In the present study, the absence of difference in global accuracy between the different angles of rotation from the upright orientation may be explained by the use of an identification task rather than a verification task.

Regarding the rotation of the symbols on the stomach, it is interesting to note that symbol recognition may have been facilitated for the 180°-rotated orientation because recognizing a 180°-rotated symbol presented on the stomach with a trunk-centred perspective is close to adopting a head-centred perspective. Indeed, adopting a head-centred perspective results in perceiving the symbol upright but mirror-reversed (except for the symmetrical symbols). As the task was not to decide whether the symbol was normal or mirror-reversed, and as there was no pair of symmetrical symbols in the lists of symbols presented to the participants, adopting a head-centred perspective may have been an easier perceptual strategy than trying to recognize rotated symbols.

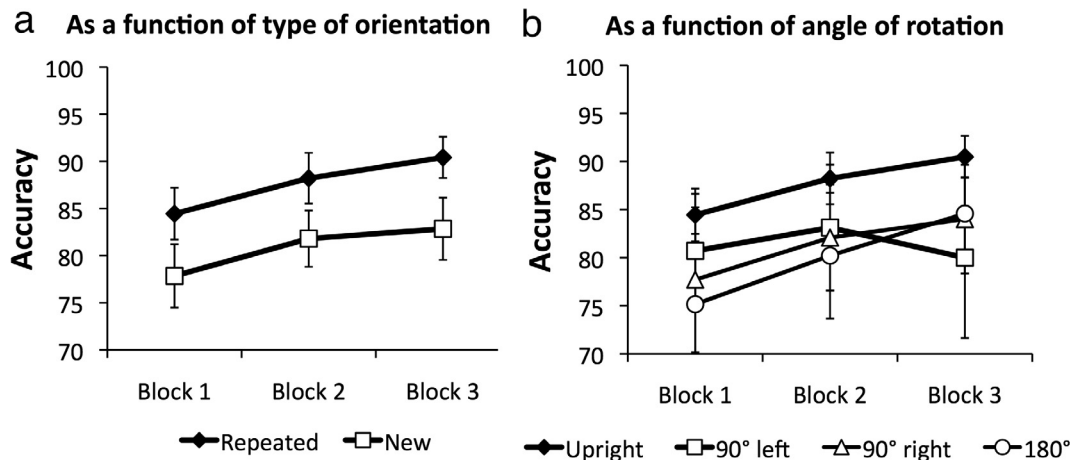


Fig. 4. Participants' accuracy in Experiment 3 as a function of (a) Type of orientation (repeated, new) and Block and (b) Angle of rotation and Block. Error bars represent the standard errors of the means.

## 5. General discussion

The present study investigated whether learning to recognize the tactile conversion of visual symbols generalizes to new stimuli and new orientations. The results show that the generalization to new stimuli depends both on the size of the stimulus set and on feature variability. Stimulus-specific learning strategies are involved when learning a limited number of stimuli whereas generalized strategies are involved when increasing the number of learned stimuli. However, in the latter case, feature variability is necessary to achieve generalization. When the learned symbols are homogeneous and composed essentially of vertical and horizontal strokes, learning does not generalize to new symbols. Our results also reveal that tactile perceptual learning generalizes to new orientations.

The present study reveals that generalization of learning depends on both the quantity and diversity of information presented during learning. Nonetheless further studies would be necessary to identify the specific factors that facilitate generalization. The most likely hypothesis is that increasing the set of learned stimuli prevents observers from involving stimulus-specific learning strategies. As a consequence, observers are forced to identify the general perceptual rules that can be applied to new stimuli. The better generalization performance reached in Experiment 2 than in Experiment 1 can also result from a difference in task difficulty. Note that if this were to be the case, this would go against the RHT of perceptual learning (Ahissar & Hochstein, 2004), which posits generalization for easy tasks and specificity for difficult tasks. In our experiments, the use of short blocks of trials and the alternation of learned and new symbols might also have contributed to generalization by preventing sensory adaptation. However, such use of short blocks and alternation were not sufficient to induce generalization in Experiment 1. Therefore it is unlikely that these characteristics of the learning protocol were solely responsible for the generalization of learning observed in Experiment 2.

The generalization to new tactile stimuli found in our study and the generalization that has been previously reported when using visual-to-auditory conversion devices (Brown & Proulx, 2013; Kim & Zatorre, 2008) can be interpreted in the light of the multisensory RHT proposed by Proulx et al. (2014). According to the classic RHT, generalization of learning benefits from broad brain activations during learning. According to the multisensory RHT, multisensory tasks produce broader brain activation than unisensory tasks in two possible ways. Cross-modal interactions might occur through direct connections between low-level unisensory areas (i.e., visual, auditory, and tactile primary cortex) or through higher-level multisensory areas where the objects are represented independently of the input sensory modality. Therefore, in the case of using a visual-to-tactile or visual-to-auditory substitution

device, direct connections can occur between tactile and visual, or between auditory and visual primary cortices. The other possibility is that the processing of tactile or auditory stimuli first progresses from tactile or auditory primary cortices to high-level object processing areas such as the lateral occipital complex (Amedi et al., 2007) and then cascade down back to visual areas. According to these two possibilities, the multisensory nature of the task produces broad brain activation and may facilitate the generalization to new stimuli.

A question remaining open is whether generalization was indeed facilitated by the multisensory nature of the task, as it corresponds to tactile recognition of visual symbols or, instead, if generalization to new stimuli is a characteristic of tactile perceptual learning itself. This is a difficult alternative to solve given that the generalization to new sets of stimuli in tactile perceptual learning has been the subject of comparatively less investigation than in visual perceptual learning. In one previous study, learning to recognize tactile patterns was found to transfer to new stimuli (Epstein, Hugues, Schneider, & Bach-y-Rita, 1989). However, the task at hand consisted in matching the tactile patterns with their visual analogue. In this case, a similar multisensory processing cannot be excluded either. In further studies, using a direct comparison of a purely tactile task with a visuo-tactile task will allow to evaluate the specific role of multisensory processing in tactile perceptual learning.

Regarding generalization to new orientations, the results obtained in our study are compatible with previous results on visual recognition of rotated alphanumeric symbols (Corballis et al., 1978; Eley, 1982), showing that alphanumeric symbols are learned through the extraction of critical features invariant to the symbol's orientation. The sequential presentation of symbols may also have reinforced orientation-invariant recognition, as spatiotemporal features rather than purely spatial ones are more likely to have been extracted here. However, the hypothesis that people are able to recognize previously learned symbols independently of their orientation seems incompatible with the influence of expertise on orientation effects. In vision, it is well known that inverting faces (i.e., 180°-rotation) dramatically decreases face recognition (Yin, 1969). This inversion effect is partly explained by the strong habit to perceive faces upright rather than rotated. An inversion effect has also been reported for unnatural objects (e.g., Greebles) that have received expertise (Gauthier, Williams, Tarr, & Tanaka, 1998). As faces, alphanumeric symbols are more frequently perceived upright than rotated. Moreover, in the tactile sensory modality, Behrmann and Ewell (2003) have shown that becoming an expert in tactile recognition induces an inversion effect. It is thus surprising that such inversion effects were not observed for the tactile recognition of visual symbols. One possibility for this discrepancy is that the duration of the present study was insufficient for participants to consolidate learning

equivalent to expertise. Another possibility is that the alternation of blocks with upright and rotated orientations, rather than a complete learning phase with upright symbols followed by test blocks with rotated symbols, prevents the facilitation effect for the upright condition. The fact that alphanumerical symbols are less complex items than faces may also have limited the involvement of specific processes such as configural processing (i.e., processing not only features, but also the relations among features), which is often characteristic of expertise (Gauthier & Tarr, 2002).

Finally, the orientation-independent perceptual learning found in our study may reflect a distal attribution process, which is an important characteristic of sensory substitution. Distal attribution corresponds to the fact that, after training to use a visual-to-tactile device, the tactile stimulation felt on the skin becomes directly attributed to external objects, as it is the case for vision. Thus distal attribution is often considered as a criterion of the involvement of visual-like processes when using a sensory substitution device. One claim made by Bach-y-Rita was that, as a consequence of distal attribution, the tactile matrix of a visual-to-tactile device can be moved from one body surface to another without loss of perceptual abilities in trained users (Bach-y-Rita & Kerckel, 2003). This claim has been confirmed by a study reporting a transfer of learning from trained to untrained body surfaces in a tactile recognition task of visual symbols (Arnold & Auvray, 2014). Distal attribution processes may also contribute here to tactile recognition of visual stimuli independently of the orientation of the tactile stimulation on the body surface.

To conclude, the present study provides new insights into the understanding of tactile perceptual learning and brain plasticity. Tactile perceptual learning appears to share some characteristics with visual perceptual learning (influence of stimulus variability and task difficulty, representation of standard features similar to visual features). This similarity may reflect the involvement of supramodal areas in both visual and tactile object recognition (Kupers & Ptito, 2011; Pascual-Leone & Hamilton, 2001). Regarding sensory substitution, the similarity in processes is compatible with the view that perception with a sensory substitution device is vertically integrated, retaining characteristics of the substituted (e.g., vision) and the substituting (e.g., audition or touch) sensory modalities (Arnold, Pesnot-Lerousseau, & Auvray, in press; Deroy & Auvray, 2012; Deroy & Auvray, 2014). There are also applied implications, in particular for the design of learning protocols for sensory substitution, suggesting that short learning sessions with high variability and diversity will facilitate generalization and will allow for a more optimal use of the device in real-life conditions. However, one remaining question is how to achieve complete generalization, that is, how to rapidly gain optimal level of performance for entirely new objects. One of the possible cues is obviously the amount of learning. Daily use of the device would allow users to become experts. Learning would also benefit from a highly diversified use of the device, with perception of different categories of complex objects and of different exemplars of the same category. Finally, an active exploration of the external world with the device, rather than a passive reception of stimuli, would also reinforce perceptual learning via the emergence of sensorimotor contingencies.

## Acknowledgment

We thank Gaëtan Parseihan and Mathieu Dubois for their help in designing the apparatus and Jess Hartcher-O'Brien and Ross Parry for their careful reading and helpful comments on previous versions of this article. This work was supported by the Labex SMART (ANR-11-LABX-65), the Mission pour l'Interdisciplinarité (CNRS, Auton, Sublima Grant, France), and the Fondation des Aveugles de Guerre. GA is funded by a grant from the European Research Council (Wearhap, FP7, N°601165).

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.visres.2017.11.007>.

## References

- Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Science*, 8, 457–464.
- Amedi, A., Stern, W. M., Camprodon, J. A., Bermpohl, F., Merabet, L., Rotman, S., & Pascual-Leone, A. (2007). Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nature Neuroscience*, 10, 687–689.
- Arno, P., Vanlierde, A., Streel, E., Wanet Defalque, M. C., Sanabria-Bohorquez, S., & Veraart, C. (2001). Auditory substitution of vision: pattern recognition by the blind. *Applied Cognitive Psychology*, 15, 509–519.
- Arnold, G., Pesnot-Lerousseau, J., & Auvray, M. (in press). Individual differences in sensory substitution. *Multisensory Research*.
- Arnold, G., & Auvray, M. (2014). Perceptual learning: Tactile letter recognition transfers across body surfaces. *Multisensory Research*, 27, 71–90.
- Arnold, G., Spence, C., & Auvray, M. (in press). A unity of the self or a multiplicity of locations? How the graphesthesia task sheds light on the role of spatial perspectives in bodily self-consciousness. *Consciousness & Cognition*.
- Arnold, G., & Siéhoff, E. (2012). Timing constraints of temporal view association in face recognition. *Vision Research*, 54, 61–67.
- Arnold, G., Spence, C., & Auvray, M. (2016). Taking someone else's spatial perspective: Natural stance or effortful decentring? *Cognition*, 148, 27–33.
- Auvray, M. (2004). *Immersion et perception spatiale: l'exemple des dispositifs de substitution sensorielle* [Immersion and spatial perception: the example of sensory substitution devices] (Unpublished doctoral dissertation). EHESS, Paris, France.
- Auvray, M., Hannonet, S., & O'Regan, J. K. (2007). Learning to perceive with a visuo-auditory substitution system: Localization and object recognition with The Voice. *Perception*, 36, 416–430.
- Auvray, M., & Myin, E. (2009). Perception with compensatory devices: From sensory substitution to sensorimotor extension. *Cognitive Science*, 33, 1036–1058.
- Bach-y-Rita, P., & Kerckel, S. W. (2003). Sensory substitution and the human-machine interface. *Trends in Cognitive Sciences*, 7, 541–546.
- Baek, A., Maes, K., Van Meel, C., & Op de Beeck, H. P. (2016). The transfer of object learning after training with multiple exemplars. *Frontiers in Psychology*, 7, 1386.
- Baek, A., Windey, L., & Op de Beeck, H. P. (2012). The transfer of object learning across exemplars and their orientation is related to perceptual similarity. *Vision Research*, 68, 40–47.
- Behrmann, M., & Ewell, C. (2003). Expertise in tactile pattern recognition. *Psychological Science*, 14, 480–492.
- Brown, D. J., & Proulx, M. J. (2013). Increased signal complexity improves the breadth of generalization in auditory perceptual learning. *Neural Plasticity*, 2013, 879047.
- Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual information processing* (pp. 75–176). New York: Academic Press.
- Corballis, M. C., Zbrodoff, N. J., Shetzer, L. I., & Butler, P. B. (1978). Decisions about identity and orientation of rotated letters and digits. *Memory and Cognition*, 6, 98–107.
- Deroy, O., & Auvray, M. (2012). Reading the world through the skin and ears: a new perspective on sensory substitution. *Frontiers in Psychology*, 3, 457.
- Deroy, O., & Auvray, M. (2014). A crossmodal perspective on sensory substitution. In S. Biggs, M. Matthen, & D. Stokes (Eds.), *Perception and its modalities* (pp. 327–349). Oxford: Oxford University Press.
- Dutilh, G., Krypotos, A. M., & Wagenmakers, E. J. (2011). Task-related versus stimulus-specific practice: A diffusion model account. *Experimental Psychology*, 58, 434–442.
- Eley, M. G. (1982). Identifying rotated letter-like symbols. *Memory and Cognition*, 10, 25–32.
- Epstein, W., Hughes, B., Schneider, S. L., & Bach-y-Rita, P. (1989). Perceptual learning of spatiotemporal events: Evidence from an unfamiliar modality. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 28–44.
- Fahle, M. (2005). Perceptual learning: Specificity versus generalization. *Current Opinion in Neurobiology*, 15, 154–160.
- Fahle, M. E., & Poggio, T. E. (2002). *Perceptual learning*. Cambridge, MA, USA: MIT Press.
- Furmanski, C. S., & Engel, S. A. (2000). Perceptual learning in object recognition: Object specificity and size invariance. *Vision Research*, 40, 473–484.
- Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 431–446.
- Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training “Greeble” experts: A framework for studying expert object recognition processes. *Vision Research*, 38, 2401–2428.
- Haggard, P., Taylor-Clarke, M., & Kennett, S. (2003). Tactile perception, cortical representation and the bodily self. *Current Biology*, 13, R170–R173.
- Harrar, V., Spence, C., & Makin, T. R. (2014). Topographic generalization of tactile perceptual learning. *Journal of Experimental Psychology: Human Perception and Performance*, 40, 15–23.
- Harris, H., Gliksberg, M., & Sagi, D. (2012). Generalized perceptual learning in the absence of sensory adaptation. *Current Biology*, 22, 1813–1817.
- Hartcher-O'Brien, J., & Auvray, M. (2014). The process of distal attribution illuminated through studies of sensory substitution. *Multisensory Research*, 27, 421–441.
- Hussain, Z., Bennet, P. J., & Sekuler, A. B. (2012). Versatile perceptual learning of

- textures after variable exposures. *Vision Research*, 61, 89–94.
- Hussain, Z., Sekuler, A. B., & Bennett, P. J. (2009). Perceptual learning modifies inversion effects for faces and textures. *Vision Research*, 49, 2273–2284.
- Jeter, P. E., Doshier, B. A., Liu, S. H., & Lu, Z. L. (2010). Specificity of perceptual learning increases with increased training. *Vision Research*, 50, 1928–1940.
- Jeter, P. E., Doshier, B. A., Petrov, A., & Lu, Z. L. (2009). Task precision at transfer determines specificity of perceptual learning. *Journal of Vision*, 9, 1–13.
- Kaczmarek, K. A., & Haase, S. J. (2003). Pattern identification and perceived stimulus quality as a function of stimulation waveform on a fingertip-scanned electrotactile display. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11, 9–16.
- Kim, J. K., & Zatorre, R. J. (2008). Generalized learning of visual-to-auditory substitution in sighted individuals. *Brain Research*, 1242, 263–275.
- Kim, J. K., & Zatorre, R. J. (2010). Can you hear shapes you touch? *Experimental Brain Research*, 202, 747–754.
- Kupers, R., & Ptito, M. (2011). Insights from darkness: What the study of blindness has taught us about brain structure and function. *Progress in Brain Research*, 192, 17–31.
- Levy-Tzedek, S., Hanassy, S., Abboud, S., Maidenbaum, S., & Amedi, A. (2012). Fast, accurate reaching movements with a visual-to-auditory sensory substitution device. *Restorative Neurology and Neuroscience*, 30, 313–323.
- Levy-Tzedek, S., Novick, I., Arbel, R., Abboud, S., Maidenbaum, S., Vaadia, E., & Amedi, A. (2012). Cross-sensory transfer of sensory-motor information: visuomotor learning affects performance on an audiomotor task, using sensory-substitution. *Scientific Reports*, 2, 949.
- Liu, Z., & Weinshall, D. (2000). Mechanisms of generalization in perceptual learning. *Vision Research*, 40, 97–109.
- Nagarajan, S. S., Blake, D. T., Wright, B. A., Byl, N., & Merzenich, M. M. (1998). Practice-related improvements in somatosensory interval discrimination are temporally specific but generalize across skin location, hemisphere, and modality. *Journal of Neuroscience*, 18, 1559–1570.
- Newell, F. N., Ernst, M. O., Tjan, B. S., & Bühlhoff, H. H. (2001). Viewpoint dependence in visual and haptic object recognition. *Psychological Science*, 12, 37–42.
- Pascual-Leone, A., & Hamilton, R. (2001). The metamodal organization of the brain. *Progress in Brain Research*, 134, 427–445.
- Pollok, B., Schnitzler, I., Mierdorf, T., Stoerig, P., & Schnitzler, A. (2005). Image-to-sound conversion: Experience-induced plasticity in auditory cortex of blindfolded adults. *Experimental Brain Research*, 167, 287–291.
- Proulx, M. J., Brown, D. J., Pasqualotto, A., & Meijer, P. (2014). Multisensory perceptual learning and sensory substitution. *Neuroscience & Biobehavioral Reviews*, 41, 16–25.
- Renier, L., Collignon, O., Poirier, C., Tranduy, D., Vanlierde, A., Bol, A., ... De Volder, A. G. (2005). Cross-modal activation of visual cortex during depth perception using auditory substitution of vision. *NeuroImage*, 26, 573–580.
- Sampaio, E., Maris, S., & Bach-y-Rita, P. (2001). Brain plasticity: 'Visual' acuity of blind persons via the tongue. *Brain Research*, 908, 204–207.
- Sathian, K., & Zangaladze, A. (1997). Tactile learning is task specific but transfers between fingers. *Perception and Psychophysics*, 59, 119–128.
- Sigman, M., & Gilbert, C. D. (2000). Learning to find a shape. *Nature Neuroscience*, 3, 264–269.
- Striem-Amit, E., Guendelman, M., & Amedi, A. (2012). 'Visual' acuity of the congenitally blind using visual-to-auditory sensory substitution. *PLoS ONE*, 7, e33136.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95, 15–48.
- Valyear, K. F., Culham, J. C., Sharif, N., Westwood, D., & Goodale, M. A. (2006). A double dissociation between sensitivity to changes in object identity and object orientation in the ventral and dorsal visual streams: a human fMRI study. *Neuropsychologia*, 44, 218–228.
- Wang, X., Zhou, Y., & Liu, Z. (2013). Transfer in motion perceptual learning depends on the difficulty of the training task. *Journal of Vision*, 13, 1–9.
- World Medical Association (1991). Declaration of Helsinki. *Law, Medicine & Health Care: A Publication of the American Society of Law & Medicine*, 19, 264–265.
- Yanagida, Y., Kakita, M., Lindeman, R. W., Kume, Y., & Tetsutani, N. (2004). Vibrotactile letter reading using a low-resolution tactor array. *Proceedings of the 12th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 400–406.
- Yi, D. J., Olson, I. R., & Chun, M. M. (2006). Shape-specific perceptual learning in a figure-ground segregation task. *Vision Research*, 46, 914–924.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81, 141–145.

# Differentiated Audio-Tactile Correspondences in Sighted and Blind Individuals

Ophelia Deroy  
University of London

Irène Fasiello, Vincent Hayward,  
and Malika Auvray  
University Pierre et Marie Curie, Sorbonne Universités

The aim of the present study is to investigate whether the crossmodal correspondence robustly documented between auditory pitch and visual elevation has analogues in the audio-tactile domain. Across 4 experiments, the compatibility effects between intuitively congruent pairs of stimuli (i.e., outward tactile movement, going from the inside of the finger toward the fingertip and increasing pitch, or inward tactile movement and decreasing pitch) and incongruent pairs stimuli (i.e., the reverse associations) were measured. Two methods were compared to assess the behavioral effects of such a correspondence: One where participants have to respond to either the auditory or tactile stimulus presented simultaneously, while ignoring the other (speeded classification task), and the other where the auditory and tactile stimuli are presented sequentially and associated to different response buttons (implicit association test). No significant compatibility effect was observed under the speeded classification task. The implicit association test revealed a significant compatibility effect. This effect was similar in the conditions where the finger was placed vertically and horizontally. However, this implicit association between pitch and tactile movements was not observed in blind participants. These results have methodological implications for the explanation and testing of crossmodal correspondences, and the origin of the widely discussed association between pitch and vertical elevation.

*Keywords:* crossmodal correspondences, audition, touch, pitch, space

We commonly use vertical elevation to refer to auditory variations: Increases in perceived pitch are experienced and reported as “rising,” while decreases in pitch are considered to “descend” or to “fall.” This spatial mapping does not strictly depend on language (Dolscheid, Hunnius, Casasanto, & Majid, 2012; Parkinson, Kohler, Sievers, & Wheatley, 2012) and can affect experience and behavior in a consistent way, from the perception of sound location (Pratt, 1930; Roffler & Butler, 1968), speeded discrimination experiments (Ben-Artzi & Marks, 1995; Bernstein & Edelstein, 1971; Melara & O’Brien, 1987), and audio-visual interactions (Parise & Spence, 2008, 2009) to the orientation of attention and the perception of ambiguous visual movements (Maeda, Kanai, & Shimojo, 2004). In most cases, the spatial mapping of pitch has been assessed in reference to the location of visual targets or the direction of visual movement (see Occelli, Spence, & Zampini,

2009, for an exception and a review). In stimulus-response compatibility effects in which participants are shown to be faster at responding to a high-pitch sound with the upper key rather than the lower one (Rusconi, Kwan, Giordano, Umiltà, & Butterworth, 2006), vision might also play a role as participants can perceive or imagine the visual organization of the two response buttons.

In the study reported here, we investigated whether the spatial mapping of pitch was intimately tied to vision. Would the perception of a change of pitch interfere with the tactile direction of movement experienced at one’s fingertips, as it does with the visual direction of movement? If the tendency to match or map pitch and gestures seem to be grounded in musical practices, the learning of a similar correspondence between pitch and passive movement, if any, would seem less easy to explain. What’s more, given the many differences in auditory, tactile, and audio-tactile perception which are introduced by visual experience (see Heller & Gentaz, 2013; Hötting & Röder, 2009 for reviews), the existence of a correspondence between pitch and tactile movement should be assessed both in blind and sighted individuals.

Besides its intrinsic interest, evidencing a correspondence between pitch and tactile movement (Deroy & Spence, 2013; Spence, 2011) could help differentiate between the different hypotheses offered to explain the pitch-elevation correspondence documented through audio-visual paradigms. According to the spatial theory of magnitude, differences in pitch are represented as “high” or “low” in an amodal mental space, like differences in brightness or numerical quantities (Rusconi et al., 2006; see also Gevers, Reynvoet, & Fias, 2003; Cohen Kadosh, Lammertyn, & Izard, 2008, for discussion). If this is indeed the case, we would expect pitch to be commonly represented in a similar spatial way, independently of

---

This article was published Online First March 7, 2016.

Ophelia Deroy, School of Advanced Study, University of London; Irène Fasiello, Vincent Hayward, and Malika Auvray, Institut des Systèmes Intelligents et de Robotique, University Pierre et Marie Curie, Sorbonne Universités.

We thank Merle Fairhurst, Jess Hartcher-O’Brien and the reviewers for their helpful comments on previous versions of the manuscript. This research was supported by the Agence Nationale de la Recherche (ANR-11-JSH2-003-1) and by the Mission pour l’Interdisciplinarité (Defisens, Supplé-Sens grant) as well as the RTS-AHRC grant (AH-L007053/1).

Correspondence concerning this article should be addressed to Ophelia Deroy, Centre for the Study of the Senses, School of Advanced Study, University of London, Senate House, Malet Street, London WC1E 7HU, United Kingdom. E-mail: [ophelia.deroy@sas.ac.uk](mailto:ophelia.deroy@sas.ac.uk)



the modality used to evidence this representation. By contrast, if the mapping between pitch and movement is primarily learned through audio-visual associations, it may differ once tested in the tactile modality. [Parise, Knorre, and Ernst \(2014\)](#) have recently showed that the natural statistics of sounds were compatible with high-pitched sounds coming from higher in space, and low-pitched sounds from lower in space. It might therefore be the case that when we hear high-pitched sounds, we associate them to the higher sources of the sounds we look at, and vice versa for low-pitched sounds. In this case, vision would be necessary for spatial mapping to be learned, and we would not expect to find a similar mapping in congenitally blind individuals. More generally, the lack of a correspondence between changes in pitch and direction of tactile movements would reinforce the idea that vision plays a role in the acquisition of the correspondence, as posited by other hypotheses (e.g., [Deroy, Crisinel, & Spence, 2013](#)). According to the transitivity hypotheses, other regular associations between sounds and visual objects can explain the pitch-elevation correspondence. For instance, an increased distance in the vertical (or even horizontal) direction correlates with a decrease in visual size: The farther or the higher an object gets, the smaller it becomes. Another environmental regularity is that the smaller an object becomes, the higher the pitch it emits ([Evans & Treisman, 2010](#); [Fernández-Prieto, Navarra, & Pons, 2015](#)). The transitivity of associative learning from higher in space to smaller in size and from smaller in size to higher in pitch could then explain how the correspondence between higher in space and higher in pitch is formed through regularities and transitive learning ([Spence & Deroy, 2012](#)). Crucially, the mediation of a change in size can only be provided by vision: A change in localization on one's body is not associated to a change in size, which seems to block the transition via "getting smaller." A specific prediction of the transitive hypothesis therefore is that the correspondence between movement and change in pitch should not exist in congenitally blind individuals.

The experiments conducted here were designed to test for the existence of an audio-tactile correspondence between pitch and tactile movements, similar to the well documented audio-visual correspondence between pitch and visual movements. Further experiments tested whether the correspondence could depend on bodily position and on visual experience. The broader goal was to understand the mapping between pitch and space.

Given the importance of attention in crossmodal congruency effects (see [Spence & Deroy, 2013](#) for review), two different methods were used to investigate the existence of an audio-tactile correspondence between tactile movements felt on one's finger and changes in pitch. These two methods focus on different attentional processing levels, one that occurs at early stages and the other at later ones, during response selection. The first method, used in many studies (e.g., [Evans & Treisman, 2010](#); [Ludwig, Adachi, & Matsuzawa, 2011](#); see [Marks, 2004](#), and [Spence, 2011](#), for reviews), draws on a speeded classification paradigm. Participants are presented with two different stimuli in different sensory modalities, and are asked to respond to one type of stimulus (for instance, high or low visual targets) while ignoring the stimulus presented in another sensory modality (for instance, high- or low-pitched sounds). The prediction is that participants are faster to respond to the attended feature when the unattended feature is congruent with it. In other words, the prediction is that people find

it harder (i.e., they are slower) to classify a target stimulus presented in one sensory modality (e.g., audition) when the distractor stimulus presented in a task-irrelevant modality (e.g., touch) varies along a dimension that shares a relation of correspondence with the target dimension (an effect which has similarities with the cross-modal Stroop test, e.g., [Cowan & Barron, 1987](#); [MacLeod, 1991](#); [Walker & Smith, 1984](#)). A second method, which has more recently been tested, builds on the implicit association test (IAT), which is meant to measure participants' automatic associations. In this task, participants are required to rapidly categorize two kinds of target stimuli whose values are associated to two response buttons, in both a congruent and an incongruent way. For instance, in the first IAT (see [Greenwald, McGhee, & Schwartz, 1998](#); see [Glashouwer, Smulders, de Jong, Roefs, & Wiers, 2013](#); [Teige-Mocigemba, Klauer, & Sherman, 2010](#), for reviews) targeting the implicit association between skin color and valence, participants would be asked to respond to either positive words or names typical of White people with the same response key (predicted to be a congruent combination) and then asked to respond either to positive words or to names typical of Black people on the same response key (predicted to be an incongruent combination). The difference in response time between the congruent and incongruent conditions reveals the strength of the implicit association. Variants of the IAT have been successfully adapted to nonsocial stimuli and used to test crossmodal correspondences between odors and touch ([Dematte, Sanabria, & Spence, 2007](#)), audition and touch ([Occelli et al., 2009](#)), taste and pitch ([Crisinel & Spence, 2010](#)), and audition and vision ([Parise & Spence, 2012](#)). The task always requires sorting four kinds of stimuli using only two response buttons, each of which refers to two of the four stimuli. For instance, for an IAT testing pitch and size, participants have to respond to high or low pitch, and large or small visual targets by pressing one of two keys, associated to either congruent (high-small, low-big) or incongruent (high-big, low-small) combinations. The prediction is that people find it easier (i.e., are overall faster) when the two stimuli that share a response are strongly associated than when they are weakly associated.

The two methods were tested in our study in order to determine whether they would give the same results for the tactile and auditory features that were selected. In the first experiment, participants had to respond to either tactile or auditory stimuli, while ignoring the simultaneous presentation of congruent or incongruent stimuli in the other sensory modality. In the second experiment, participants had to respond to either a tactile or an auditory feature presented one after the other in blocks and with congruent or incongruent responses that were associated to the same response key. In a third experiment, we tested whether the audio-tactile correspondence would depend on the hand position, and thereby if it took place within a hand-centered frame of reference. To do so, the participants completed the experiment with the arm placed in two directions: in the vertical and horizontal planes. Finally, in order to investigate the role of vision, a fourth experiment looked at the existence of audio-tactile crossmodal correspondences in early and late blind participants.

### Experiment 1: Speeded Classification Protocol

A first experiment looked for a crossmodal congruency effect between pitch and tactile movements by means of a speeded

classification paradigm. Participants had to categorize auditory stimuli while ignoring simultaneously presented tactile stimuli and vice versa.

## Method

**Participants.** Sixteen sighted participants (11 men and five women) took part in this experiment.<sup>1</sup> Their ages ranged from 19 to 40 years ( $M = 29$  years). All of the participants were naive to the purpose of the experiment and reported normal auditory and tactile perception. They received €6 in return for their participation. The experiment took approximately 45 min to complete. The four experiments reported in this article were performed in accordance with the ethical standards laid down in the 1991 Declaration of Helsinki.

**Apparatus and stimuli.** The tactile stimuli were delivered by a distributed tactile stimulator (Wang, & Hayward, 2010) from Tactile Lab, Inc. (Latero). This device induces lateral tractions on the skin's surface by means of a matrix of miniature piezoelectric bending actuators (64 actuators in an  $8 \times 8$  matrix). The active area of the tactile display is  $1.0 \text{ cm}^2$ . The Latero was controlled by a 3.40GHz i7 2600 PC equipped with a dedicated Ethernet interface running under the User Datagram Protocol. The software ran in the Xenomai Real-Time Framework on Linux (Ubuntu), guaranteeing a stable rate for the control loop at high frequency. The refresh rate of the display was 1 kHz. Both the tactile and auditory stimuli were run from the same computer.

For the tactile stimuli, a virtual bar was simulated by the tactile display by locally inducing oscillatory strains in the skin of the fingertip at a rate of 80 Hz. The simulated bar was orthogonal to the finger's principal axis (see Figure 1). To minimize positional cues at the start of the stimulation, the tactile stimuli started from the central rows of the 9.6-mm long matrix, and then moved either toward the tip of the finger or toward the palm. From now on, we refer to these movements as outward and inward movements respectively. The stimulus duration was 250 ms, thus giving rise to a bar moving at a speed of 19.2 mm/s. The neutral stimulation consisted in a random activation of the four central rows of the tactile matrix.

The auditory stimuli were presented through a loudspeaker which was located just behind the tactile display in order to ensure a spatial coincidence between tactile and auditory stimuli. The stimuli consisted in three pure tones of linearly increasing (700 to

1,200 Hz), linearly decreasing (700 to 200 Hz), or constant (700 Hz) pitch. The duration of each of the auditory stimuli was 250 ms.

**Procedure.** The experiment was run in a dark, anechoic chamber. The participants sat in front of a table which also supported the tactile device and the loudspeaker, located approximately 25 cm from the participant. Participants completed two main sessions, one where they had to respond to the tactile stimuli, and one where they had to respond to the auditory ones. The order of presentation of these sessions was counterbalanced across participants. Each session was made of a block of training trials, followed by a unimodal block of trials, then a bimodal block of trials.

During the training block, participants learned the mapping between the stimuli and the relevant response keys. They were required to press the left or right button of the mouse in order to listen to the associated auditory stimulus (during the auditory session) or to experience the tactile stimulation (during the tactile session). In all blocks, the mouse buttons were therefore orthogonal to the direction of tactile movement. During this training, only unimodal stimuli were presented, and no verbal label was assigned to them. The stimuli were prompted by the participant: For instance, the participant heard a sound increasing in pitch when pressing the right button of the mouse and a sound decreasing in pitch when pressing the left button of the mouse (and vice versa for half of the participants). Similarly, during the tactile session, the participant had to place either his index or major finger on the tactile display (choosing the one he felt the most comfortable with). He would experience an inward or outward tactile direction of movement when pressing the right or the left key of the mouse. Participants were encouraged to listen to the auditory and experience the tactile stimuli until they were sure that they had learnt the mapping. They were first given 20 practice trials. Those participants who felt they had not learnt the stimulus-response mapping were given additional trials until they reported the mapping without errors.

Following a successful practice block, participants underwent one unimodal block of trials. As in the training block, only two kinds of tactile and auditory stimuli were used: inward or outward tactile stimulation (during the tactile session) and sounds of increasing or decreasing pitch (during the auditory session). Each stimulus was presented 10 times, resulting in a total of 20 unimodal trials for each modality (and 40 for the two modalities).

The participants' task was to respond to the direction of tactile movement or changes in pitch by pressing the corresponding right and left buttons of a mouse (the association between stimuli and buttons were counterbalanced across participants). Participants were instructed to respond as accurately and as quickly as possible, and they could make their responses at any time up to 2,500 ms after stimulus onset. If the participant responded before the stimulus ended, this response would

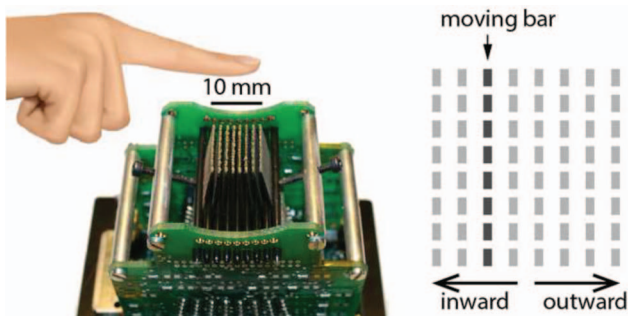


Figure 1. Tactile stimulator device and representation of the tactile stimulus. See the online article for the color version of this figure.

<sup>1</sup> Note that two participants performed at chance level in the unimodal tactile baseline condition, which means that they were not able, owing to the novelty of the task, to discriminate the tactile stimuli. These participants did not complete the rest of the experiment and were not counted among the participants reported here. Similar inability to discriminate tactile stimuli were found in the subsequent experiments; thus, two participants were excluded from Experiment 2, five participants were excluded from Experiment 3, and none were excluded from Experiment 4.

be logged as such but the stimulation was still completed. Participants were given auditory feedback regarding the correctness of their responses: Whenever they provided an incorrect answer, they heard the word “incorrect.” Both the verbal instructions and feedback were prerecorded by the experimenter. There was then a 500-ms intertrial interval; that is, between the offset of the stimulus or the offset of the feedback in case of incorrect trials, and the beginning of the next trial. Participants were given periodic breaks after every 20 trials.

In the bimodal blocks of trials, participants experienced 12 possible pairs of stimuli. When they had to respond to the tactile stimuli, the target tactile stimulus could be either inward or outward, and could be accompanied by one of three possible concurrent auditory stimuli: either increasing, decreasing, or neutral change in frequency (i.e., six possible pairings, see Table 1). In the auditory session, that is, when they had to respond to the auditory stimuli, the target auditory stimulus could be either of increasing or decreasing frequency, and could be accompanied by one of three possible concurrent tactile stimuli: inward, outward, or neutral (i.e., without direction; six other pairings, see Table 1). Each of the bimodal blocks of trials consisted of 30 repetitions of the six possible combinations. Participants thus completed a total of 360 trials over the tactile and auditory sessions.

Participants had their fingertip resting on the tactile display and listened to the sounds presented by the loudspeakers. On each test trial, participants experienced the tactile and the auditory stimuli simultaneously. They were required to pay attention to only one of them while ignoring the other. They were instructed to classify the stimuli as accurately and as quickly as possible by pressing one of the two keys of the mouse. As in the unimodal blocks, participants could make their response at any time up to 2,500 ms after stimulus onset. Verbal feedback was provided to them in cases of incorrect responses and there was then a 500-ms break between the offset of the stimulus and the beginning of the next trial. Participants were given periodic breaks after every 20 trials.

## Results

Trials in which participants failed to provide a response before the trial was terminated (overall less than 0.5% of trials) were not included in the data analyses. The primary dependent measure was the reaction time (RT), measured from stimulus onset, on correct trials only. Note that we also recorded the accuracy of the responses and 92.2% of the responses were correct. An analysis of variance (ANOVA) was conducted on the mean RTs for correct answers given by each participant with two factors: Modality (tactile vs. auditory) and Congruency (unimodal, neutral, congruent, and incongruent). The analysis revealed a significant main effect of Modality,  $F(1, 15) = 119.95, p < .00001$ , a significant

main effect of Congruency,  $F(3, 45) = 9.17, p < .0001$ , and a significant interaction between Modality and Congruency,  $F(3, 45) = 4.82, p < .006$ . A Duncan post hoc test on Modality  $\times$  Congruency revealed that, when the target stimulus was tactile, there were significant differences between the unimodal condition and the three other conditions ( $p < .001$ ); no other significant difference appeared. When the target stimulus was auditory there was no significant difference between the levels of the congruency factor. Participants were faster overall in the auditory conditions ( $M = 54.87 \pm SD$  of 7.63 ms) compared with the tactile ( $M = 83.06 \pm 10.47$  ms) condition.

To summarize the results, a significant difference between the unimodal and the bimodal sessions was evidenced, but the difference between the congruent and the incongruent conditions did not reach significance either for the tactile or the auditory condition. In other words, this experiment failed to reveal a crossmodal congruency effect (see Figure 2).

## Interim Discussion

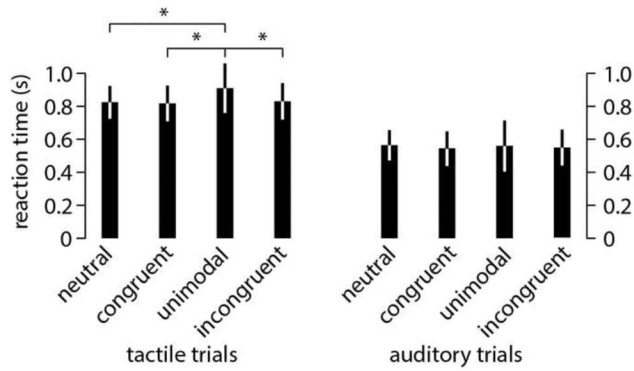
The results did not reveal a congruency effect between direction of pitch and direction of tactile movement although the method used here was similar to other studies which obtained positive congruency effects between direction of pitch and visual movement (Evans & Treisman, 2010). One problem with the joint presentation of stimuli though comes from a lack of control of selective attention. Indeed, nothing guarantees that participants' attention was focused on the targeted sensory modality only, and that the results do not reflect a failure of selective attention (e.g., Parise & Spence, 2012, for a discussion; see also Melara & O'Brien, 1987). As no significant difference between tones of increasing or decreasing frequency was observed when participants were asked to respond to touch, we can infer that participants did not merely base their answers on the perceived direction of pitch. The second experiment, resting on the successive presentation of stimuli and relying on their association with different response buttons, offers an alternative way to control for the role of attention.

## Experiment 2: Implicit Association Task

The second experiment investigated the crossmodal correspondence between auditory changes in pitch and tactile direction of movement by means of an implicit association test. Participants were presented with either a sound or a tactile stimulus and they had to categorize the stimulation (ascending vs. descending pitch, inward vs. outward tactile direction). They first had to learn the association between the four possible stimuli and the two response buttons that could be either congruent (i.e., the same button for

Table 1  
*The Twelve Possible Combinations of Auditory and Tactile Stimuli Presented in the Bimodal Blocks*

Response to simultaneous presentations	Tactile stimulation		Auditory stimulation	
	Outward	Inward	Rising	Descending
Congruent	Rising pitch	Descending pitch	Outward tactile movement	Inward tactile movement
Incongruent	Descending pitch	Rising pitch	Inward tactile movement	Outward tactile movement
Neutral	Neutral pitch change	Neutral pitch change	Random tactile stimulation	Outward tactile movement



*Figure 2.* Participants' mean reaction times for the two conditions of stimulus presentation: tactile and auditory; for the four congruency conditions: congruent (upward tactile and rising pitch; downward tactile and descending pitch), unimodal (tactile only; auditory only), incongruent (upward tactile and descending pitch; downward tactile and rising pitch), and neutral (upward tactile and neutral pitch; downward tactile and neutral pitch in the tactile block; neutral tactile and rising pitch; neutral tactile and descending pitch in the auditory block). Error bars represent the standard errors of the means.

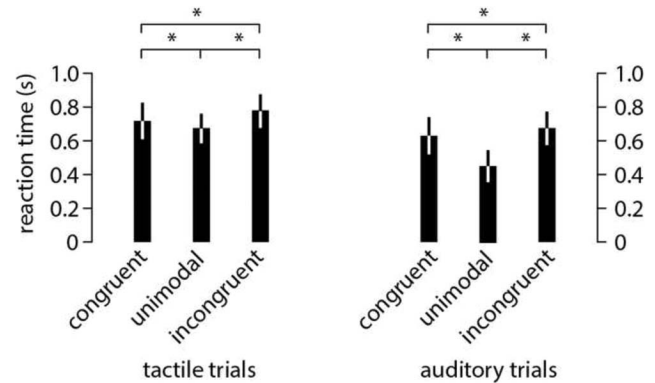
increasing pitch and outward direction of movement, the other one for decreasing pitch and inward tactile direction of movement) or incongruent (the same button for increasing pitch and inward tactile direction of movement; the other one for decreasing pitch and outward tactile direction of movement). If there is a cross-modal correspondence between pitch and tactile direction of movement, participants are expected to be more accurate and faster in the congruent condition than in the incongruent one.

## Method

**Participants.** Fourteen sighted participants (seven men and seven women) took part in this experiment. Their ages ranged from 20 to 32 years ( $M = 26$  years). All the participants were naive to the purpose of the experiment and reported normal auditory and tactile perception. They received €12 in return for their participation in this experiment which took approximately 90 min to complete.

**Procedure.** The materials (i.e., the tactile device and auditory display) were similar to those reported for Experiment 1. The arm on which the tactile stimuli were presented remained in a horizontal position. The tactile and auditory stimuli had the same durations as in Experiment 1. The tasks and procedure were different: this time an implicit association test was used (see below). Participants went through a training session and then completed a unimodal session made up of one tactile block and one auditory block (whose order was counterbalanced across participants). After the unimodal session, participants completed a bimodal session.

The training was similar to Experiment 1 in that participants had to learn the mapping between the stimuli and the relevant response keys, without explicit verbal labels. In the unimodal tactile block, the participants experienced inward and outward tactile stimulation. Their task was to respond to each stimulus by pressing one of the two mouse buttons. In the unimodal auditory block, the task was the same but participants experienced sounds increasing and decreasing in frequency. Each type of stimulus was repeated 10



*Figure 3.* Participants' mean reaction times for the two conditions of stimulus presentation: tactile and auditory; for the three congruency conditions: congruent (upward tactile and rising pitch; downward tactile and descending pitch associated to the same response key), unimodal (upward tactile and downward tactile associated to two response keys; rising pitch and descending pitch associated to two response keys), and incongruent (upward tactile and descending pitch; downward tactile and rising pitch associated to the same response key). Error bars represent the standard errors of the means.

times, for a total of 40 unimodal trials across the tactile and auditory modalities.

In the bimodal sessions, participants first learned the association between the four possible stimuli and the two response buttons on the computer mouse. The associations were planned so as to be either congruent (i.e., increasing pitch and outward tactile direction of movement on the same button, decreasing pitch and inward tactile direction of movement on the other button) or incongruent (increasing pitch and inward tactile direction of movement on the same button, decreasing pitch and outward direction of movement on the other). These labels were not provided to the participants, who were just trained to respond to the stimuli. Participants completed the two possible associations in separate blocks, with half of the cohort starting with the congruent condition and half with the incongruent one.

On each trial, participants experienced either a tactile or an auditory stimulus and were required to identify them as accurately and as quickly as possible. When the stimulus was auditory, they had to press one of the two mouse buttons depending on whether the sound was increasing or decreasing in pitch. When the stimulus was tactile, they had to press one of the mouse buttons depending on whether the direction of movement was going inward or outward. Each session consisted of 30 repetitions of each of the two tactile stimuli and auditory stimuli presented in a random order. Participants completed 120 trials for each congruent and incongruent session (i.e., 240 in total).

## Results

Trials in which participants failed to provide a response before the trial was terminated (overall less than 0.4% of trials) were not included in the data analyses. The primary dependent measure was the RT, measured from stimulus onset, on correct trials only (Figure 3). The accuracy of the responses was also recorded and reached 93.88% of correct answers. An ANOVA conducted on the

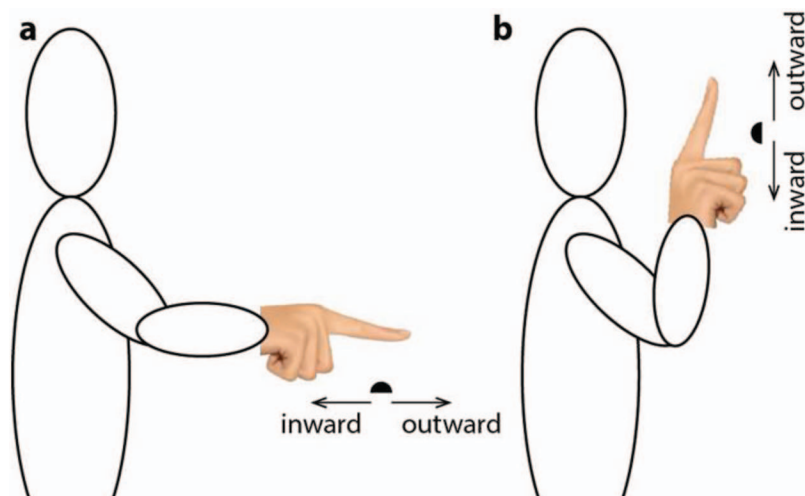


Figure 4. The two arm positions used in Experiments 3 and 4: vertical and horizontal. See the online article for the color version of this figure.

mean RTs on the correct answers for each participant with two factors: Modality (tactile vs. auditory) and Congruency (unimodal, congruent, and incongruent) revealed a significant main effect of Modality,  $F(1, 13) = 8.62, p < .00001$ , a significant main effect of Congruency,  $F(2, 26) = 62.90, p < .00001$ , and a significant interaction between Modality and Congruency,  $F(2, 26) = 35.62, p < .00001$ . A Duncan post hoc test on Modality  $\times$  Congruency revealed that, except between auditory unimodal and tactile incongruent, all the other interactions were significant (all  $ps < .01$ ). In particular, there was a significant difference for tactile stimuli between the congruent and incongruent conditions ( $p = .0002$ ); and for auditory stimuli as well ( $p = .0013$ ), with participants being faster in the congruent conditions than in the incongruent ones. Overall, participants were faster in the auditory ( $68.84 \pm 7.98$  ms) than in the tactile ( $M = 85.84 \pm 10.89$  ms) condition. They were faster in the unimodal ( $67.02 \pm 7.15$  ms) followed by the congruent ( $80.35 \pm 11.04$  ms) and then the incongruent condition ( $86.90 \pm 10.40$  ms).

### Interim Discussion

In this variant of the IAT, participants were required to focus on one stimulus at a time, while the association was only at the level of response buttons. The results revealed a significant congruency effect between changes in pitch and changes in direction of tactile movement. Participants were for instance faster in responding to a tone increasing in pitch when the response button was shared with the outward tactile movement than when it was shared with the inward tactile movement. This result suggests that an association between tactile movement and direction of pitch operates at the implicit level. Besides the role of attention, this protocol avoids the comparison or direct mapping required by the speeded classification task used in the first experiment. When presented together, the assumption is that a decrease in pitch should correspond to a downward tactile movement, and an increase in pitch to an upward tactile movement. As the primary sensory organs of vision and audition are rigidly attached to the head together with the vestibular

system, and as the head is typically maintained in a vertical position, “upward” and “downward” for vision and audition seem to correspond to the vertical—that is, to the head orientation. In contrast, touch is distributed throughout the entire body: As tactile receptive surfaces are permanently reoriented in space, it is not clear how the direction of tactile movement is encoded. To develop this hypothesis further, we conducted a third experiment, using the successful method used in the second experiment.

### Experiment 3: Implicit Association Test: Horizontal Versus Vertical Arm Position

The third experiment investigated the influence of arm position on the crossmodal correspondence between change in auditory pitch and change in tactile direction of movement. The position of the arm on which participants received the tactile stimuli was placed, in two separate conditions, either vertically or horizontally (see Figure 4). If a greater correspondence effect is obtained in the vertical condition, this would suggest that the most natural spatial mapping of pitch corresponds to the vertical plane, as it is the case for the audio-visual correspondence between pitch and vertical elevation in visual space (e.g., Chiou & Rich, 2012; Rusconi et al., 2006, for comparisons between different spatial mappings of pitch). On the other hand, if a similar effect is obtained across the two conditions, the congruency would be effective both in the horizontal and vertical plane. This could mean that, with respect to touch, the bodily reference frame (the hand) is determinant for the correspondence and that the tip of the finger corresponds to “up,” independently of the actual hand position. Alternatively, it could also be the case that there are two different correspondences at stake: one between pitch and the vertical plane, and the other between pitch and the horizontal plane.

### Method

**Participants.** Fourteen sighted participants (one man and 13 women) took part in this experiment. Their ages ranged from 19 to

34 years ( $M = 25$  years). All of the participants were naive to the purpose of the experiment and reported normal auditory and tactile perception. They received €12 in return for their participation in this experiment which took approximately 90 min to complete.

**Materials and procedure.** The materials and procedure were similar to those reported for Experiment 2 with the following difference: The participants' arm position was horizontal in one session and vertical in the other (see Figure 4). The order of these two sessions was counterbalanced across participants. The only difference was the number of trials in the bimodal session which was the same for all the participants, but divided into the two conditions of arm movements. More precisely, the participants completed 15 repetitions of each of the two tactile stimuli and auditory stimuli for each of the congruency conditions (congruent and incongruent) and for each of the arm positions (vertical and horizontal), that is, 240 trials in total.

## Results

Trials in which participants failed to provide a response before the trial was terminated (overall less than 0.53% of trials) were not included in the data analyses. The accuracy of the responses reached 93.10% of correct answers. An ANOVA was conducted on the mean RTs for correct trials for each participant, with three factors: Modality (tactile vs. auditory), Arm Position (horizontal vs. vertical), and Congruency (unimodal, congruent, and incongruent). The ANOVA revealed a significant main effect of Modality,  $F(1, 13) = 41.59, p < .0001$ , a significant main effect of Congruency,  $F(2, 26) = 58.38, p < .00001$ , and no effect of Arm Position,  $F(1, 13) = .06, p = .81$ . There was a significant interaction between Modality and Congruency,  $F(2, 26) = 40.36, p < .00001$ . There was no other significant interaction (all  $F$ s  $< 1$ ). A Duncan post hoc test on Modality  $\times$  Congruency revealed that except between tactile unimodal and auditory incongruent all the other interactions were significant (all  $p$ s  $< .05$ ). In particular, there was a significant difference for responses to tactile stimuli between the congruent and incongruent conditions ( $p < .001$ ), with participants being faster in the congruent condition. Similarly, there was a significant difference for responses to auditory stimuli between the congruent and incongruent conditions ( $p = .001$ ) with participants being faster in the congruent condition (Figure 5). Thus, this experiment revealed a crossmodal congruency effect for the two sensory modalities and the two arm positions.

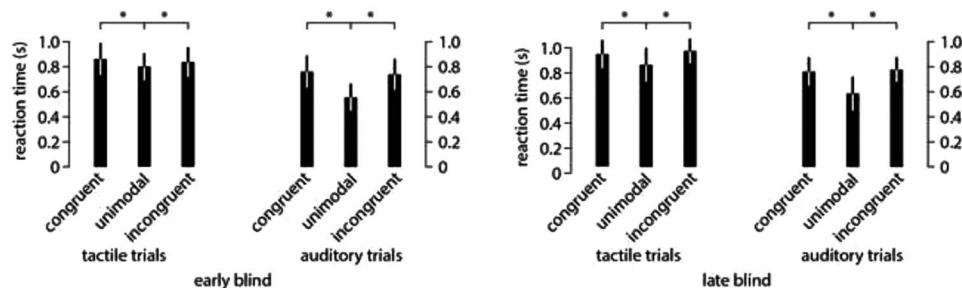


Figure 5. Participants' mean reaction times for the two conditions of stimulus presentation: tactile and auditory, and for the three congruency conditions: congruent, unimodal, and incongruent (as in Experiment 2). Error bars represent the standard errors of the means. Note that, as there was no effect of arm position, in the graph, horizontal and vertical arm positions were collapsed.

Overall, the participants were faster in the auditory ( $72.19 \pm 7.16$  ms) than in the tactile ( $M = 90.09 \pm 12.70$  ms) condition. They were also faster in the unimodal ( $70.74 \pm 7.96$  ms) followed by the congruent ( $84.66 \pm 8.98$  ms) and then the incongruent condition ( $90.75 \pm 12.21$  ms).

## Interim Discussion

The third experiment confirmed the existence of an implicit association between changes in pitch and direction of tactile movement. More importantly, it revealed that this association occurs in two different arm positions. Thus, a rising pitch is associated to a tactile movement going toward the fingertip, and a descending pitch to a tactile movement going inward toward the palm, independently of the hand being oriented upward or horizontally. Results are consistent with the correspondence occurring in a hand-centered frame of reference.

## Experiment 4: Horizontal Versus Vertical Arm Position in Early and Late Blind People

The fourth experiment investigated the role of prior visual experience on the crossmodal correspondence between auditory pitch and tactile direction of movement, while still investigating the role of arm position in the correspondence. Early blind and late blind participants completed the experiment as described in Experiment 3.

## Method

**Participants.** Twenty-four blind participants took part in this experiment. Twelve of them were late blind (four men and eight women, becoming blind after the age of five), eight of them were congenitally blinds (five men and three women), and four of them were early blind, that is, they became blind between 1 and 3 years of age (one man and three women). Note that in the results, as there were no significant differences between the eight congenitally blind and the four early blind, their results were subsequently regrouped. All blind participants were completely blind or had light perception without shape perception and used a cane or a seeing-eye dog for walking. For all blind participants, blindness resulted from genetic disease, retinitis, an accident or cancer and was not associated to any additional neurological impairment. All

participants were naive to the purpose of the experiment and reported normal auditory and tactile perception. They received €25 in return for their participation in this experiment which took approximately 180 min to complete.

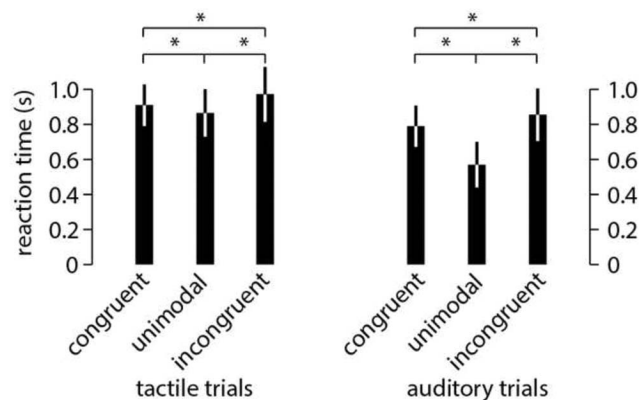
**Materials and procedure.** The materials and procedure were the same as Experiments 2 and 3.

## Results

Trials in which participants failed to provide a response before the trial was terminated (overall less than 0.56% of trials) were not included in the data analyses. The accuracy of the responses reached 93.92% of correct answers. An ANOVA was conducted on mean RTs, for the correct answers given by each participant only, with four factors: Onset of Blindness (early vs. late), Modality (tactile vs. auditory), Arm Position (horizontal vs. vertical), and Congruency (unimodal, congruent, and congruent). Onset of Blindness was a between-participants factor whereas Modality, Arm Position, and Congruency were within-participants factors. The ANOVA revealed a significant main effect of Modality,  $F(1, 22) = 92.72, p < .00001$ , a significant main effect of Congruency,  $F(2, 44) = 60.12, p < .00001$ , no effect of Onset of Blindness,  $F(1, 22) = .04, p = .32$ , and no effect of Arm Position,  $F(1, 22) = .10, p = .75$ . There was a significant interaction between Modality and Congruency,  $F(2, 44) = 20.65, p < .00001$ . There was no other significant interactions, except between Arm Position, Modality, and Congruency,  $F(2, 44) = 4.05, p = .024$ . A Duncan post hoc test on Modality  $\times$  Congruency revealed that when the target stimulus was tactile there were significant differences between the unimodal condition and the two other conditions (all  $ps < .001$ ) but not between the congruent and incongruent conditions ( $p = .98$ ). Similarly, when the target stimulus was auditory, there were significant differences between the unimodal condition and the two other conditions (all  $ps < .001$ ) but not between the congruent and incongruent conditions ( $p = .90$ ). Participants were overall faster in the auditory ( $M = 61.50 \pm 6.15$  ms) than in the tactile ( $M = 76.02 \pm SD$  of 9.34 ms) condition. The participants were faster in the unimodal ( $61.36 \pm 7.41$  ms) than both in the congruent ( $73.23 \pm 8.96$  ms) and the incongruent ( $73.11 \pm 7.76$  ms) conditions. In summary, although the results revealed a significant effect of congruency there were no significant differences between the congruent and incongruent condition in either of the two modalities. Thus there was no crossmodal congruency effect for both early and late blind participants (see Figure 6). Correlations were also conducted and no correlation was found between the amount of previous visual experience and the size of the effect (congruent–incongruent) in the blind participant group,  $r(22) = .32, p = .12$ . No correlation was found for the participants of the late blind group only,  $r(10) = .13, p = .69$ .

## Interim Discussion

The performance of blind participants did not show a congruency effect between the direction of tactile movement and changes in pitch, whereas this effect was evidenced in sighted participants in Experiments 2 and 3, which used the same methodology. What's more, there was no effect related to the age of onset. This failure cannot be attributed to a failure of attention, as the experiment relied on stimuli being successively presented. This differ-



*Figure 6.* Participants' mean reaction times for the two conditions of stimulus presentation: tactile and auditory; for the three congruency conditions: congruent, unimodal, and incongruent (as in Experiment 2 and 3); and for the two onsets of blindness: early and late. Error bars represent the standard errors of the means. Note that, as there was no effect of arm position, in the graph, horizontal and vertical arm positions were collapsed.

ence is not linked to arm position either, as blind participants did not exhibit a congruency effect in either the horizontal or vertical arm positions. Vision seems therefore to be necessary to mediate the correspondence between direction of pitch and direction of tactile movement.

## General Discussion and Conclusions

The present study reveals the existence of a congruency effect between auditory changes in pitch and changes in direction of movement in touch. The study also provides further evidence of the usefulness of IAT protocols when testing crossmodal correspondences. Interestingly though, whereas effects of audio-visual congruence between pitch and visual elevation have been successfully tested in speeded classification tasks (Evans & Treisman, 2010; Marks, 2004), no significant effect was observed in our audio-tactile version. This lack of an effect in the speeded classification task, but not in the IAT task, might be explained by the attentional demand needed to ignore the random presentation of one kind of stimuli. It could also come from a lack of strategic information about the informative character of the paired stimuli.

A difference between the present protocol and previous ones comes from the fact that the participants in our study were not explicitly instructed about the informative character of the other cue (e.g., they were told that a high-pitched sound is more likely to be correlated with a high visual target, Chiou & Rich, 2012) or at least trained by using the predicates “high” and “low” to describe the sounds (Evans & Treisman, 2010). Here, we purposely provided only neutral information about these characteristics, to avoid biasing participants on the classification of the tactile stimuli. The lack of a label for the “rising–descending” character of the sounds and tactile stimuli might have been responsible for the lack of effect in the cognitively demanding speeded classification task. As was suggested by various authors (see Spence & Deroy, 2013, for a review), the effects of crossmodal correspondences on behavior, when tested in multisensory settings, might not be automatic and depend therefore on an explicit strategic

representation. In future research, one way to test for the influence of explicit representations on performance would be to compare the results of the first experiment with and without specific “leading” instructions. This can be done for instance by saying to participants that the sounds and/or tactile stimuli are rising versus descending, or by designing a test where they are told that the correspondence between sounds and direction of inward and outward movements is informative.

The results obtained in the IAT experiment extend the results obtained by [Occelli et al. \(2009\)](#) regarding the correspondence between pitch height and tactile elevation. They reveal the existence of a congruency effect between the dynamic changes in pitch and the direction of tactile movement. The results obtained in our study also go further in terms of explaining the origin of the spatial mapping of pitch which underlies a well-documented series of effects (e.g., [Evans & Treisman, 2010](#); [Patching & Quinlan, 2002](#); [Rusconi et al., 2006](#)). They challenge the idea of a general, amodal spatial representation of pitch shared across all modalities, along with other spatial representations of magnitudes (e.g., [Lourenco & Longo, 2011](#)). This does not mean that tasks involving linguistic labels might not show a general spatial representation of pitch (e.g., see [Bottini, Crepaldi, Casasanto, Crollen, & Collignon \(2015\)](#) for the role of language on the spatial representations of time in the blind). The present results strongly suggest the need to investigate whether other spatial mappings of magnitudes, documented with vision, also hold for touch.

The present results are difficult to reconcile with the semantic hypothesis developed by [Martino and Marks \(1999\)](#) and by [Walker and Walker \(2012\)](#) to explain the presence of crossmodal correspondences. According to this hypothesis, correspondences come from the fact that different stimuli evoke similar concepts, or fall within the same part of a conceptual dichotomy (e.g., passive/active; pleasant/unpleasant). Blind participants, who refer to changes in pitch in terms of “rising” and “descending,” have no reason not to fail to apply similar concepts to the two stimuli presented here and to map these directions of pitch and direction of tactile movement. Further evidence of the role of semantic mapping is needed, and could be obtained by testing this audio-tactile correspondence in sighted participants whose language does not describe pitch in spatial terms (see for instance [Parkinson et al., 2012](#), for a test with the audiovisual correspondence). The difference between crossmodal correspondence and semantic associations suggested here adds to the results obtained by [Sadaghiani et al. \(2009\)](#) who found that crossmodal associations between pitch and height were mediated by different neurological processes than metaphorical descriptors such as “high pitch” and “low pitch.”

Importantly, the absence of an effect in blind participants suggests that visual mediation plays a necessary role in the acquisition and persistence of the pitch-elevation correspondence. Our results add evidence to the role of vision in space-magnitude interactions (on the absence of spatial-numerical association of response codes [SNARC] effect in early blind individuals, see [Crollen, Dormal, Seron, Lepore, & Collignon, 2013](#)). In line with the present suggestion, [Alink, Euler, Kriegeskorte, Singer, and Kohler \(2012\)](#) have also found that auditory directional information was transferred to high-level visual cortex in healthy adults. Making sense of the role of vision in the pitch-direction correspondence opens several hypotheses. As mentioned in the introduction, vision might be necessary to associate higher pitched sounds to the higher

locations in space in which they most regularly originate ([Parise et al., 2014](#)). In this sense, the extraction of information from auditory scene statistics would require a crossmodal collaboration. Alternatively, vision might be necessary to mediate a transitive association between objects getting higher and therefore smaller, and emitting higher pitch as they get smaller. It is only in vision, and not in touch, that moving objects are associated to a change in objects’ size, and change in size is in turn associated to a change in pitch (see [Eitan, Schupak, Gotler, & Marks, 2011](#) for a recent study). However, if the absence of association mediated by vision might then explain why congenitally blind participants did not acquire the correspondence between tactile direction and change in pitch, the absence of an association in late blind also suggests a role of visual exposure in the persistence of the association. Due to the presence of visual imagery in the late blind, it could be interesting to test the transitive hypothesis by training them with the two associations (direction of tactile movement – change in [imagined] visual size) and (change in [imagined] visual size – change in pitch) to investigate whether this might lead to different results. It could also be the case that the mental notion of verticality for blind people is determined by the direction of the gravity load due to hand-held objects whereas in sighted people, it is driven by the visual aspect of standing structures. As such they do not refer to identical notions.

Here the findings should be related to [Mossbridge, Grabowecy, and Suzuki \(2011\)](#), who measured the influence of auditory frequency changes on visuo-spatial attention. The influence disappeared when participants’ head axis was tilted by 90°, and no longer aligned with the body axis, which led the author to conclude that “because this cross-modal cueing is dependent on the alignment of head-centered and environmental axes, it may develop through associative learning during waking upright experience” ([Mossbridge et al., 2011](#), p. 133). As was noted earlier, hands occupy many different positions in space. They are not constrained by vestibular inputs that enable us to orient our heads vertically when walking, looking, and listening ([Pozzo, Berthoz, & Lefort, 1990](#); [Berthoz, 1991](#)). Hands and fingertips, instead, are completely mobile; they operate independently from gravity and are subject only to proprioceptive relationships. Whereas there is good evidence that, in sighted and in late blind individuals, touch is encoded in a visual frame of reference (e.g., [Azañón, Longo, Soto-Faraco, & Haggard, 2010](#); [Harrar & Harris, 2010](#)), this fact seems insufficient to explain the presence of an effect in sighted individuals when the hand is horizontal, and the absence of an effect in both hand positions for late blinds. As such, the fact that sighted people seem to associate a movement going from the inner palm to the tip of the finger to an upward movement (and at least to a rising pitch) for two different hand positions opens an interesting question regarding the implicit directionality of hands in the body schema: Whereas vision modulates implicit hand maps when it comes to size and shape ([Longo, 2014](#)), and the crossing of two hands has well documented effects on crossmodal tasks ([Shore, Spry, & Spence, 2002](#)), little has been done to test whether a default position or direction of one single hand is used in speeded responses. A further study could be performed to test whether these correspondences are maintained when the hands point in various directions and are located in various relationships with the body. Alternatively, given that the weighing of skin-based, anatomical coordinates and external spatiotopic coordinates is task-



dependent (Badde, Röder, & Heed, 2015), it cannot be excluded that the reference frame observed here was triggered by the task, rather than a default one.

To conclude, the present results reveal differences in the audio-visual congruency exhibited by blind and sighted people. Other differences have recently been shown in the classical sound symbolic Bouba-Kiki task (Fryer, Freeman, & Pring, 2014) as well as in SNARC effects (Crollen et al., 2013). The present study should help to develop user-centered methodologies, accounting for a diversity of potential users (e.g., Deroy & Auvray, 2012). Several devices dedicated to assisting blind individuals (e.g., the vOICe, Meijer, 1992; Vibe, Hanneton, Auvray, & Durette, 2010) or navigation in the absence of vision (e.g., the Flybar, see Pedley & Harper, 1959) have been designed by using the correspondence between direction of pitch and direction of movement (higher–higher; lower–lower) but without testing whether this correspondence, intuitive for the designers and well-documented in audio-visual contexts for sighted individuals, was also relevant for blind people. Here, the relevance of this coding is questioned given that this correspondence does not show in blind people. A better understanding of visual impairment shows that there is not a single abstract profile, but rather a continuum of cases ranging across individuals and situations: defects or accidents, total or partial, stable or progressive, constant or variable depending on the situation, with or without other deficits. The IAT test of crossmodal correspondences might be an interesting tool to pretest the relevance of certain codings or combinations of auditory and tactile cues, to tailor them more specifically to their users. Cross-Modal Correspondences Enhance Performance on a Colour-to-Sound Sensory Substitution Device, Multisensory Research). Our study also raise interesting questions as to the multiplicity of spatial frames in which changes of pitch sounds can be mapped.

## References

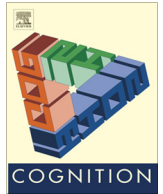
- Alink, A., Euler, F., Kriegeskorte, N., Singer, W., & Kohler, A. (2012). Auditory motion direction encoding in auditory cortex and high-level visual cortex. *Human Brain Mapping, 33*, 969–978. <http://dx.doi.org/10.1002/hbm.21263>
- Azañón, E., Longo, M. R., Soto-Faraco, S., & Haggard, P. (2010). The posterior parietal cortex remaps touch into external space. *Current Biology, 20*, 1304–1309. <http://dx.doi.org/10.1016/j.cub.2010.05.063>
- Badde, S., Röder, B., & Heed, T. (2015). Flexibly weighted integration of tactile reference frames. *Neuropsychologia, 70*, 367–374. <http://dx.doi.org/10.1016/j.neuropsychologia.2014.10.001>
- Ben-Artzi, E., & Marks, L. E. (1995). Visual-auditory interaction in speeded classification: Role of stimulus difference. *Perception & Psychophysics, 57*, 1151–1162. <http://dx.doi.org/10.3758/BF03208371>
- Bernstein, I. H., & Edelman, B. A. (1971). Effects of some variations in auditory input upon visual choice reaction time. *Journal of Experimental Psychology, 87*, 241–247. <http://dx.doi.org/10.1037/h0030524>
- Berthoz, A. (1991). Reference frames for the perception and control of movement. In J. Paillard (Ed.), *Brain and space* (pp. 81–111). New York, NY: Oxford University Press.
- Bottini, R., Crepaldi, D., Casasanto, D., Crollen, V., & Collignon, O. (2015). Space and time in the sighted and blind. *Cognition, 141*, 67–72. <http://dx.doi.org/10.1016/j.cognition.2015.04.004>
- Chiou, R., & Rich, A. N. (2012). Cross-modality correspondence between pitch and spatial location modulates attentional orienting. *Perception, 41*, 339–353. <http://dx.doi.org/10.1068/p7161>
- Cohen Kadosh, R., Lammertyn, J., & Izard, V. (2008). Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Progress in Neurobiology, 84*, 132–147. <http://dx.doi.org/10.1016/j.pneurobio.2007.11.001>
- Cowan, N., & Barron, A. (1987). Cross-modal, auditory-visual Stroop interference and possible implications for speech memory. *Perception & Psychophysics, 41*, 393–401. <http://dx.doi.org/10.3758/BF03203031>
- Crisinel, A. S., & Spence, C. (2010). A sweet sound? Food names reveal implicit associations between taste and pitch. *Perception, 39*, 417–425. <http://dx.doi.org/10.1068/p6574>
- Crollen, V., Dormal, G., Seron, X., Lepore, F., & Collignon, O. (2013). Embodied numbers: The role of vision in the development of number-space interactions. *Cortex, 49*, 276–283. <http://dx.doi.org/10.1016/j.cortex.2011.11.006>
- Demattè, M. L., Sanabria, D., & Spence, C. (2007). Olfactory-tactile compatibility effects demonstrated using a variation of the Implicit Association Test. *Acta Psychologica, 124*, 332–343. <http://dx.doi.org/10.1016/j.actpsy.2006.04.001>
- Deroy, O., & Auvray, M. (2012). Reading the world through the skin and ears: A new perspective on sensory substitution. *Frontiers in Psychology, 3*, 457. <http://dx.doi.org/10.3389/fpsyg.2012.00457>
- Deroy, O., Crisinel, A. S., & Spence, C. (2013). Crossmodal correspondences between odors and contingent features: Odors, musical notes, and geometrical shapes. *Psychonomic Bulletin & Review, 20*, 878–896. <http://dx.doi.org/10.3758/s13423-013-0397-0>
- Deroy, O., & Spence, C. (2013). Why we are not all synesthetes (not even weakly so). *Psychonomic Bulletin & Review, 20*, 643–664. <http://dx.doi.org/10.3758/s13423-013-0387-2>
- Dolscheid, S., Hunnius, S., Casasanto, D., & Majid, A. (2012). The sound of thickness: Prelinguistic infants' associations of space and pitch. In *CogSci 2012: The 34th annual meeting of the Cognitive Science Society* (pp. 306–311). Austin, TX: Cognitive Science Society.
- Eitan, Z., Schupak, A., Gotler, A., & Marks, L. E. (2011). Lower pitch is larger, yet falling pitches shrink: Interaction of pitch change and size change in speeded discrimination. *Proceedings of Fechner Day, 27*, 81–88.
- Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision, 10*, 1–12.
- Fernández-Prieto, I., Navarra, J., & Pons, F. (2015). How big is this sound? Crossmodal association between pitch and size in infants. *Infant Behavior and Development, 38*, 77–81. <http://dx.doi.org/10.1016/j.infbeh.2014.12.008>
- Fryer, L., Freeman, J., & Pring, L. (2014). Touching words is not enough: How visual experience influences haptic-auditory associations in the “Bouba-Kiki” effect. *Cognition, 132*, 164–173. <http://dx.doi.org/10.1016/j.cognition.2014.03.015>
- Gevers, W., Reynvoet, B., & Fias, W. (2003). The mental representation of ordinal sequences is spatially organized. *Cognition, 87*, B87–B95. [http://dx.doi.org/10.1016/S0010-0277\(02\)00234-2](http://dx.doi.org/10.1016/S0010-0277(02)00234-2)
- Glashouwer, K. A., Smulders, F. T., de Jong, P. J., Roefs, A., & Wiers, R. W. (2013). Measuring automatic associations: Validation of algorithms for the Implicit Association Test (IAT) in a laboratory setting. *Journal of Behavior Therapy and Experimental Psychiatry, 44*, 105–113. <http://dx.doi.org/10.1016/j.jbtep.2012.07.015>
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*, 1464–1480. <http://dx.doi.org/10.1037/0022-3514.74.6.1464>
- Hanneton, S., Auvray, M., & Durette, B. (2010). The vibe: A versatile vision-to-audition sensory substitution device. *Applied Bionics and Biomechanics, 7*, 269–276. <http://dx.doi.org/10.1155/2010/282341>
- Harrar, V., & Harris, L. R. (2010). Touch used to guide action is partially coded in a visual reference frame. *Experimental Brain Research, 203*, 615–620. <http://dx.doi.org/10.1007/s00221-010-2252-0>

- Heller, M., & Gentaz, E. (2013). *The psychology of touch and blindness*. New York, NY: Taylor & Francis.
- Hötting, K., & Röder, B. (2009). Auditory and auditory-tactile processing in congenitally blind humans. *Hearing Research*, 258, 165–174. <http://dx.doi.org/10.1016/j.heares.2009.07.012>
- Longo, M. R. (2014). The effects of immediate vision on implicit hand maps. *Experimental Brain Research*, 232, 1241–1247. <http://dx.doi.org/10.1007/s00221-014-3840-1>
- Lourenco, S. F., & Longo, M. R. (2011). Origins and development of generalized magnitude representation. In S. Dehaene & E. M. Brannon (Eds.), *Space, time and number in the brain: Searching for the foundations of mathematical thought* (pp. 225–244). London, United Kingdom: Academic. <http://dx.doi.org/10.1016/B978-0-12-385948-8.00015-3>
- Ludwig, V. U., Adachi, I., & Matsuzawa, T. (2011). Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (*Pan troglodytes*) and humans. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 20661–20665. <http://dx.doi.org/10.1073/pnas.1112605108>
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109, 163–203. <http://dx.doi.org/10.1037/0033-2909.109.2.163>
- Maeda, F., Kanai, R., & Shimojo, S. (2004). Changing pitch induced visual motion illusion. *Current Biology*, 14, R990–R991. <http://dx.doi.org/10.1016/j.cub.2004.11.018>
- Marks, L. E. (2004). Cross-modal interactions in speeded classification. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of multisensory processes* (pp. 85–105). Cambridge, MA: MIT Press.
- Martino, G., & Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: Tests of the semantic coding hypothesis. *Perception*, 28, 903–923. <http://dx.doi.org/10.1068/p2866>
- Meijer, P. B. L. (1992). An experimental system for auditory image representations. *IEEE Transactions on Bio-Medical Engineering*, 39, 112–121. <http://dx.doi.org/10.1109/10.121642>
- Melara, R. D., & O'Brien, T. P. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General*, 116, 323–336. <http://dx.doi.org/10.1037/0096-3445.116.4.323>
- Mossbridge, J. A., Grabowecky, M., & Suzuki, S. (2011). Changes in auditory frequency guide visual-spatial attention. *Cognition*, 121, 133–139. <http://dx.doi.org/10.1016/j.cognition.2011.06.003>
- Occelli, V., Spence, C., & Zampini, M. (2009). Compatibility effects between sound frequency and tactile elevation. *Neuroreport*, 20, 793–797. <http://dx.doi.org/10.1097/WNR.0b013e32832b8069>
- Parise, C. V., Knorre, K., & Ernst, M. O. (2014). Natural auditory scene statistics shapes human spatial hearing. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 6104–6108. <http://dx.doi.org/10.1073/pnas.1322705111>
- Parise, C., & Spence, C. (2008). Synesthetic congruency modulates the temporal ventriloquism effect. *Neuroscience Letters*, 442, 257–261. <http://dx.doi.org/10.1016/j.neulet.2008.07.010>
- Parise, C. V., & Spence, C. (2009). “When birds of a feather flock together”: Synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PLoS ONE*, 4, e5664. <http://dx.doi.org/10.1371/journal.pone.0005664>
- Parise, C. V., & Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: A study using the implicit association test. *Experimental Brain Research*, 220, 319–333. <http://dx.doi.org/10.1007/s00221-012-3140-6>
- Parkinson, C., Kohler, P. J., Sievers, B., & Wheatley, T. (2012). Associations between auditory pitch and visual elevation do not depend on language: Evidence from a remote population. *Perception-London*, 41, 854–861. <http://dx.doi.org/10.1068/p7225>
- Patching, G. R., & Quinlan, P. T. (2002). Garner and congruence effects in the speeded classification of bimodal signals. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 755–775.
- Pedley, P. E., & Harper, R. S. (1959). Pitch and the vertical localization of sound. *The American Journal of Psychology*, 72, 447–449. <http://dx.doi.org/10.2307/1420051>
- Pozzo, T., Berthoz, A., & Lefort, L. (1990). Head stabilization during various locomotor tasks in humans. I. Normal subjects. *Experimental Brain Research*, 82, 97–106. <http://dx.doi.org/10.1007/BF00230842>
- Pratt, C. C. (1930). The spatial character of high and low tones. *Journal of Experimental Psychology*, 13, 278–285. <http://dx.doi.org/10.1037/h0072651>
- Roffler, S. K., & Butler, R. A. (1968). Localization of tonal stimuli in the vertical plane. *The Journal of the Acoustical Society of America*, 43, 1260–1266. <http://dx.doi.org/10.1121/1.1910977>
- Rusconi, E., Kwan, B., Giordano, B. L., Umiltà, C., & Butterworth, B. (2006). Spatial representation of pitch height: The SMARC effect. *Cognition*, 99, 113–129. <http://dx.doi.org/10.1016/j.cognition.2005.01.004>
- Sadaghiani, S., Maier, J. X., & Noppeney, U. (2009). Natural, metaphorical, and linguistic auditory direction signals have distinct influences on visual motion processing. *The Journal of Neuroscience*, 29, 6490–6499. <http://dx.doi.org/10.1523/JNEUROSCI.5437-08.2009>
- Shore, D. I., Spry, E., & Spence, C. (2002). Confusing the mind by crossing the hands. *Cognitive Brain Research*, 14, 153–163. [http://dx.doi.org/10.1016/S0926-6410\(02\)00070-8](http://dx.doi.org/10.1016/S0926-6410(02)00070-8)
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception & Psychophysics*, 73, 971–995. <http://dx.doi.org/10.3758/s13414-010-0073-7>
- Spence, C., & Deroy, O. (2012). Crossmodal correspondences: Innate or learned? *i-Perception*, 3, 316–318. <http://dx.doi.org/10.1068/i0526ic>
- Spence, C., & Deroy, O. (2013). How automatic are crossmodal correspondences? *Consciousness and Cognition*, 22, 245–260. <http://dx.doi.org/10.1016/j.concog.2012.12.006>
- Teige-Mocigemba, S., Klauer, K. C., & Sherman, J. W. (2010). A practical guide to Implicit Association Tests and related tasks. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 117–139). New York, NY: Guilford Press.
- Walker, P., & Smith, S. (1984). Stroop interference based on the synaesthetic qualities of auditory pitch. *Perception*, 13, 75–81. <http://dx.doi.org/10.1068/p130075>
- Walker, P., & Walker, L. (2012). Size-brightness correspondence: Cross-talk and congruity among dimensions of connotative meaning. *Attention, Perception & Psychophysics*, 74, 1226–1240. <http://dx.doi.org/10.3758/s13414-012-0297-9>
- Wang, Q., & Hayward, V. (2010). Biomechanically optimized distributed tactile transducer based on lateral skin deformation. *The International Journal of Robotics Research*, 29, 323–335. <http://dx.doi.org/10.1177/0278364909345289>

Received August 14, 2014

Revision received August 26, 2015

Accepted August 31, 2015 ■



## Short Communication

## Taking someone else's spatial perspective: Natural stance or effortful decentring?

Gabriel Arnold<sup>a,b,\*</sup>, Charles Spence<sup>c</sup>, Malika Auvray<sup>a</sup><sup>a</sup> Institut des Systèmes Intelligents et de Robotique, CNRS UMR 7722, Université Pierre et Marie Curie, Paris, France<sup>b</sup> Institut Jean Nicod (ENS – EHESS – CNRS), Département d'Etudes Cognitives, Ecole Normale Supérieure, PSL Research University, Paris, France<sup>c</sup> Department of Experimental Psychology, Crossmodal Research Laboratory, University of Oxford, Oxford, United Kingdom

## ARTICLE INFO

## Article history:

Received 9 February 2015

Revised 7 December 2015

Accepted 12 December 2015

Available online 23 December 2015

## Keywords:

Perspective taking

Sense of self

Tactile perception

Visuo-spatial abilities

Individual differences

## ABSTRACT

When perceiving stimuli, self-centred and decentred perspectives can be adopted. In the present study, we investigate whether perceivers have a natural perspective that constrains their spatial perception, with some people perceiving better with self-centred than decentred perspectives and vice versa for other people. We used a recognition task of tactile ambiguous letters (b, d, p, and q) presented on the stomach, for which three perspectives can be adopted (trunk-centred, head-centred, and decentred). At first, the participants were free to adopt any perspective they wanted. Then, either the same or a different perspective was imposed on them. Without constraints, 80% of the participants adopted a self-centred perspective (50% trunk-centred, 30% head-centred) and 20% a decentred one. The perspective adopted freely appears to be natural as recognition performance decreases with a different perspective and returns to its previous high level with the same perspective. Thus, to perceive space, some perceivers adopt naturally a perspective centred on themselves whereas others take naturally others' perspective.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

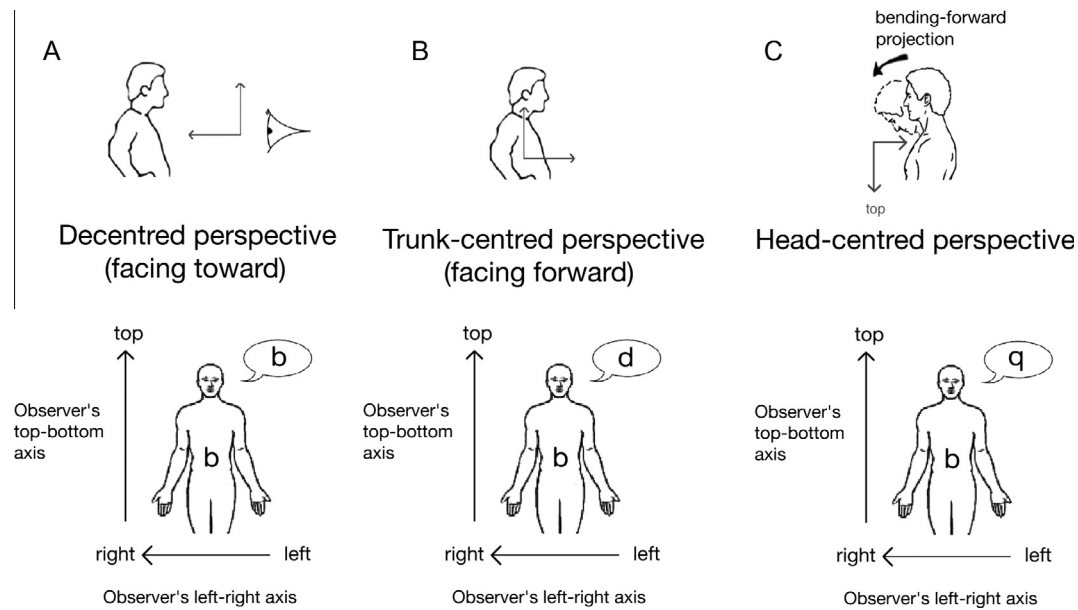
Perceivers can adopt different spatial perspectives that are either centred on their own location (i.e., self-centred) or else on a different location (i.e., decentred). On the one hand, self-centred perspectives underlie self-consciousness by binding together the multisensory experiences and the physical body (Ferrè, Lopez, & Haggard, 2014; Vogeley & Fink, 2003). Consequently, self-centred perspectives are often seen as having some sort of primacy in terms of spatial cognition (e.g., Epley, Morewedge, & Keysar, 2004). On the other hand, the ability to adopt the perspective of others is crucial when it comes to communicating and interacting with them (Schober, 1993). Decentred perspectives can be spontaneously adopted in collaborative situations (Duran, Dale, & Kreuz, 2011) and even in situations where communication is not required (Thirioux, Jorland, Bret, Tramus, & Berthoz, 2009; Tversky & Hard, 2009). In addition, as a consequence of certain personality traits – for instance, being dominated or dominant (see Galinsky, Magee, Inesi, & Gruenfeld, 2006) – some people adopt the perspective of others whereas other people

adopt their own perspective. One important question is whether people have a natural stance to adopt either self-centred or decentred perspectives. In the present study, we target this question using the tactile ambiguous symbol recognition task.

Tactile perception is interesting for the investigation of spatial perspectives because self-centred (e.g., perception from our body) and decentred perspectives (e.g., perception from outside the body) conflict with each other. Moreover, more than one self-centred perspective exists: the perspective can be centred either on the stimulated surface or on a central body part (e.g., the head; Harrar & Harris, 2010; Ho & Spence, 2007). In this sense, the recognition of ambiguous tactile symbols displayed on the body surface such as the letters b, d, p, and q (Ferrè et al., 2014; Natsoulas & Dubanovski, 1964; Parsons & Shimojo, 1987; Sekiyama, 1991; for a review, see Arnold, Spence, & Auvray, submitted for publication) provides an excellent paradigm with which to investigate the perspectives that are naturally adopted by perceivers. The same perceived stimulation can be interpreted as corresponding to different symbols, as a function of the perspective that is taken when interpreting the stimulation. For example, when the letter b is drawn on a participant's stomach (from the viewpoint of the experimenter located in front of them), three different perspectives can be adopted (see Fig. 1): a decentred perspective oriented toward the participant's stomach (response b); a trunk-centred perspective oriented forward the participant (response d); a

\* Corresponding author at: Institut des Systèmes Intelligents et de Robotique, Université Pierre et Marie Curie, 4 place Jussieu, 75005 Paris, France.

E-mail address: [arnold@isir.upmc.fr](mailto:arnold@isir.upmc.fr) (G. Arnold).



**Fig. 1.** Illustration of the three possible perspectives that participants can adopt when interpreting ambiguous symbols displayed on the body surface. In this figure, the lowercase letter “b” is drawn on the participant’s stomach from the experimenter’s perspective. Top row: the spatial perspective that can be inferred from the participant’s responses. Bottom row: the different responses reported by participants. (A) Perception of the letter “b”, resulting from the adoption of a decentred perspective whose origin is located in front of the participant. (B) Perception of the mirror-reversed letter “d”, resulting from a trunk-centred perspective. The horizontal and vertical axes of the letter are assigned congruently to the participant’s trunk. (C) Perception of the 180°-rotated letter “q”, resulting from a bending-forward head-centred perspective.

head-centred perspective, as if the head was bending forward to “see” the tactile stimulation (response q).

Important inter-individual differences have been observed in the recognition of ambiguous tactile symbols with preferences to adopt one of the three possible perspectives (Sekiyama, 1991). However, one important question that has not been directly addressed is whether perceivers have a natural perspective that constrains their spatial perception. Do some perceivers perceive better from a self-centred perspective whereas other perceivers perceive better from a decentred perspective? The aim of the present study was therefore to investigate whether individual preferences for self-centred vs. decentred perspectives reflect the natural perspectives that people adopt. Tactile symbols were presented on the stomach, for which three different perspectives can be adopted (see Fig. 1). In order not to risk biasing the participants toward the experimenter’s perspective, symbols were not drawn manually by the experimenter (which was the case in all previous studies) but by means of a matrix of vibrators.

To test the *natural perspective* hypothesis, the instructions given to the participants were varied in terms of the perspective that was to be adopted. First, the participants were free to adopt any perspective that they wanted, thus allowing us to evaluate their baseline preferences. Second, different perspectives were imposed on the participants. These corresponded either to the same or to a different perspective than the one that they had adopted freely. These imposed perspectives allowed for the evaluation of any cost associated with adopting an unnatural perspective. If the perspective that is adopted freely is natural, then imposing a different perspective should produce a cost in terms of recognition performance. Some perceivers should perform better with self-centred than decentred perspectives and vice versa for other perceivers. On the other hand, if participants are not constrained by a natural perspective, one possibility is that imposing a different perspective should not induce a cost. However, as decentred perspectives are more demanding than self-centred perspectives (Epley et al., 2004; Natsoulas, 1966), another

possibility is that all of the perceivers would perform better with a self-centred than with a decentred perspective, independently of the perspective adopted freely. Finally, in order to evaluate whether the cost of adopting an unnatural perspective is simply explained by changes in perspectival instructions or by the difficulty that is associated with disengaging from a perspectival choice, we evaluated whether performance would improve when the participants returned to the natural perspective after adopting an unnatural one. If the cost of adopting a different perspective is explained simply by changes in instruction or the difficulty that is associated with disengaging from a perspectival choice, returning to the natural perspective adopted freely should not increase performance.

In addition, we evaluated whether the ability to adopt an unnatural perspective would be influenced by visuo-spatial abilities and by the natural perspective. We thus compared the cost of adopting an unnatural perspective in those participants who adopted the trunk-centred, head-centred, and decentred perspective. However, only the two perspectives for which the vertical axis is not reversed (i.e., the trunk-centred and decentred) were imposed. The decentred perspective was imposed on participants who freely adopted the trunk-centred perspective and vice versa for the decentred participants. For the head-centred participants, the trunk-centred perspective was imposed for one half and the decentred for the other half. The head-centred perspective was not imposed because the top–bottom axis is less prone to confusion than the left–right axis. Left–right confusion occurs when a self-centred or a decentred perspective is imposed on participants (Natsoulas, 1966). However, vertical confusion is less frequent because the vertical assignment is influenced, on the one hand, by both the external environment (i.e., gravity) and the orientation of the egocentric top–bottom axis (i.e., the head–foot axis; Oldfield & Phillips, 1983), and, on the other, by the orientation of the head with a head-centred perspective (Sekiyama, 1991). Moreover, when the vertical axis is reversed, consequently to the adoption of a head-centred perspective, there is only one possible left–right assignment.

## 2. Methods

### 2.1. Participants

Eighty participants completed the experiment (44 females; mean age = 26.6 years, range = 19–47), including participants who adopted freely the trunk-centred ( $N=20$ ), the decentred ( $N=20$ ), and the head-centred ( $N=40$ ) perspectives. The head-centred participants were divided into two groups ( $N=20$ ), one adopting the trunk-centred perspective and the other the decentred perspective. In addition to the 80 participants who performed the present experiment, a further 170 participants performed the first session and were then included in other studies (see Participants in the [Supplementary Materials](#) for details concerning the classification of the participants in the different groups). The participants provided informed consent and received payment for taking part in the study. The experiment was performed in accordance with the ethical standards laid down in the [Declaration of Helsinki \(1991\)](#).

### 2.2. Apparatus

The tactile stimuli were presented by means of 9 rectangular vibrators (Haptuator Mark II, Tactile Labs, Montreal, Canada) arranged in a 3-by-3 array with a centre-to-centre spacing of 5 cm (see [Fig. 2a](#)). A nine-channel amplifier drove each vibrator independently at a frequency of 250-Hz. The vibrator array was placed on the participant's stomach symmetrically to their body mid-sagittal line. Only one layer of clothing was allowed between the skin and the vibrators and the participants individually selected the intensity of each vibrator by means of a method of adjustment. The participants wore noise-reducing headphones with a noise reduction rating of 30 dB, in order to mask any sounds made by the vibrators.

### 2.3. Stimuli

The lower-case letters b, d, p, and q were presented. The tracing of these letters consisted of a sequence of 8 vibrotactile stimuli mapping the trajectory of vibrations as if the letters were traced beginning from the stem (see [Fig. 2b](#)). The same order of strokes was used for each letter, instead of respecting the conventions of normal manual writing, because the order in which the various strokes are made in normal manual drawing can itself provide a cue to letter recognition ([Parkinson & Khurana, 2007](#)). Note that it could be argued that tracing the letters from the stem end could have biased the participants' responses toward the letter "b", given that only this letter is written in this way. However, each letter was reported equally often by participants (24.8% of trials for the b,

24.9% for the d, 25.5% for the p, and 24.8% for the q;  $F(3, 237) < 1, ns$ ). The duration of each vibration was 250 ms with no interval between consecutive vibrations, resulting in a total duration of 2-s for the presentation of each letter.

### 2.4. Procedure

The experiment was composed of three sessions of the letter recognition task, followed by the completion of the Mental Rotation Test (MRT; [Vandenberg & Kuse, 1978](#)) and the Object Perspective Taking Test (OPTT; [Hegarty & Waller, 2004](#)). In Session 1, the participants were free to adopt any perspective that they wanted to recognize letters. In Sessions 2 and 3, they were instructed to adopt a specific perspective. In one of these two sessions, the imposed perspective was the same as in Session 1. In the other session, the imposed perspective was different. Half of participants performed Session 2 with the same perspective as in Session 1 and performed Session 3 with the different one. The other half did the opposite.

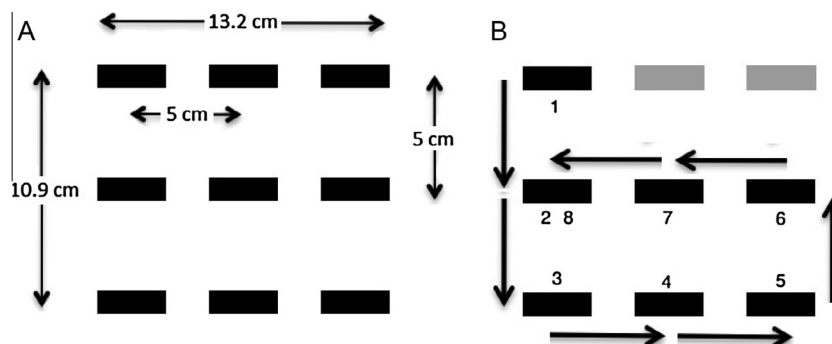
Participants gave their responses by pressing the corresponding key on the computer keyboard with the index finger of their preferred hand. The participants were instructed to keep their head oriented upright during the letter tracing, so that they do not see their stomach. The participants were able to give their responses at any time from the onset of the first vibration and up to 3000 ms after the end of the last vibration. At the end of each trial, there was an interval of 3000 ms before the start of the next trial.

Each of the three sessions was composed of 3 blocks of 16 trials (4 presentations of each of the 4 letters). At the beginning of Session 1, the participants performed a practice block with one presentation of each letter. At the beginning of Sessions 2 and 3, the practice block was composed of 3 presentations of each letter. During the practice blocks, feedback was presented indicating that the participant has given a response that was either correct or incorrect. Feedback was not presented during the test blocks but the participants were informed of their percentages of correct responses and mean response times at the end of each block.

## 3. Results

### 3.1. Proportions of each perspective adopted freely and consistency in the perspective adopted

Across all participants ( $N=250$ ; 150 females; mean age = 25.2 years, range = 18–47) who completed the task under free instructions, 49.6% adopted the trunk-centred, 29.2% the head-centred, and 21.2% the decentred perspectives. Around 4 out of 5 participants thus adopted a self-centred perspective (centred either on



**Fig. 2.** (A) Schematic figure illustrating the 3 × 3 array of rectangular vibrators. (B) The sequence of 8 vibrations for drawing the letter b. The duration of each vibration was 250 ms, without intervals between each vibration.

their head or on their trunk), whereas only 1 out of 5 adopted a decentred perspective.

For the 80 participants who performed the entire experiment, the consistency in the perspective adopted was high: 82% in Block 1, 92% in Block 2, and 96% in Block 3 (see the [Supplementary Materials](#) for more details on consistency). The consistency was not significantly different for those participants adopting the trunk-centred (88.4%), the head-centred (93.3%), and the decentred (88.8%) perspectives ( $F(2,77) = 1.00$ ;  $p > .250$ ). There was no significant difference between the 3 groups with respect to response times (RTs) during Session 1 ( $F(2,77) < 1$ ; *ns*). The three perspectives were thus not more or less demanding when participants were free to adopt their natural perspective.

### 3.2. Cost of perspective change

In order to evaluate the cost of adopting an unnatural perspective, ANOVAs were conducted on accuracy and RTs with Block (1-1, 1-2, 1-3, 2-1, 2-2, 2-3, 3-1, 3-2, 3-3; the first number indicates the Session and the second the Block) as a within-participant factors and the Order of imposed perspectives (same-different, different-same) as a between-participants factor. With accuracy, there was a significant main effect of Block ( $F(8,624) = 9.45$ ;  $p < .001$ ;  $\eta^2 = .108$ ) and a significant interaction between Block and Order of imposed perspectives ( $F(8,624) = 7.77$ ;  $p < .001$ ;  $\eta^2 = .091$ ). The cost of perspective change was evaluated by comparing recognition performance in blocks 1-3 and 2-1 for those participants who adopted the different perspective in Session 2 (different-same order), and in blocks 2-3 and 3-1 for those participants adopting the different perspective in Session 3 (same-different order). There was a significant cost associated with participants adopting the different perspective in Session 2 (a decrease of 9.3 percentage points in accuracy;  $F(1,78) = 25.44$ ;  $p < .001$ ;  $\eta^2 = .246$ ) and for participants adopting the different perspective in Session 3 (a decrease of 7.2 percentage points;  $F(1,78) = 11.65$ ;  $p = .001$ ;  $\eta^2 = .130$ ) (see [Fig. 3a](#)).

There were also training effects within sessions, with the participants making fewer errors after consecutive blocks with the same perspective (significant for both groups in Session 1, all  $ps < .001$ ; approaching significance for the different perspective in Session 2,  $p = .054$ , and the different perspective in Session 3,  $p = .080$ ). In order to make sure that the cost of adopting a different perspective was not merely explained by a lack of training, we compared recognition performance in Blocks 1-1 and 2-1, in Blocks 1-2 and 2-2, and in Blocks 1-3 and 2-3, for those participants who adopted the different perspective in Session 2. There was a significant cost in Block 2-2 (a decrease of 5.9 percentage points in accuracy;  $F(1,78) = 6.80$ ;  $p < .05$ ;  $\eta^2 = .075$ ) and 2-3 (a decrease of 5.6 percentage points;  $F(1,78) = 10.65$ ;  $p < .01$ ;  $\eta^2 = .130$ ), but not in Block 1 2-1 (an increase of 3.4 percentage points;  $F(1,78) < 1$ ; *ns*). However, performance in Block 1-1 was low, probably because the participants were not yet familiar with the task. Except for the first block, performance was thus better for the natural than for the unnatural perspective, even when the participants received the same amount of training in each perspective. Note that we did not make these comparisons for those participants adopting the different perspective in Session 3 because they kept their natural perspective in Session 2 and received thus more training with the natural than the unnatural perspective. However, the cost was not significantly different for participants adopting the unnatural perspective in Session 2 (decrease of 9.4 percentage points in accuracy) and in Session 3 (decrease of 7.2 percentage points;  $t(78) < 1$ ; *ns*).

Importantly, a significant improvement in accuracy was also observed for participants returning to their perspective in Session 3 (an increase of 5.3 percentage points in accuracy;  $F(1,78) = 6.36$ ;  $p = .014$ ;  $\eta^2 = .075$ ). This improvement indicates that the

decrease in recognition performance with the different perspective was neither produced by a change in perspectival instructions, nor by the difficulty associated with disengaging from a perspectival choice, but by the difficulty that the participants experienced when trying to adapt to an unnatural perspective.

With respect to the RT data, there was a significant main effect of Block ( $F(8,624) = 46.89$ ;  $p < .001$ ;  $\eta^2 = .375$ ) and a significant interaction between Block and Order of imposed perspectives ( $F(8,624) = 43.78$ ;  $p < .001$ ;  $\eta^2 = .359$ ) ([Fig. 3b](#)). A significant cost of adopting an unnatural perspective was observed in Session 2 (an increase of 436 ms in RT;  $F(1,78) = 45.64$ ;  $p < .001$ ;  $\eta^2 = .369$ ) and in Session 3 (an increase of 768 ms;  $F(1,78) = 70.48$ ;  $p < .001$ ;  $\eta^2 = .475$ ). RTs were also significantly shorter for participants returning to their natural perspective in Session 3 (a decrease of 441 ms;  $F(1,78) = 23.22$ ;  $p < .001$ ;  $\eta^2 = .229$ ).

### 3.3. Cost as a function of the natural perspective

We evaluated whether the cost of adopting an unnatural perspective varied with the natural perspective of the participant (see [Fig. 4](#)). With respect to the accuracy data, the cost was significantly different from zero for the trunk-centred participants when adopting the decentred perspective (a decrease of 10.6 percentage points in accuracy;  $Z = 5.23$ ;  $p < .001$ ), the head-centred participants when adopting the trunk-centred perspective (a decrease of 7.2 percentage points;  $Z = 2.33$ ;  $p = .010$ ), and the head-centred participants when adopting the decentred perspective (a decrease of 12.8 percentage points;  $Z = 3.15$ ;  $p = .001$ ). For the decentred participants adopting the trunk-centred perspective, the difference approached significance (a decrease of 2.5 percentage points;  $Z = 1.41$ ;  $p = .080$ ). A one-way ANOVA revealed that the cost was significantly smaller for the decentred than for the self-centred participants ( $F(1,76) = 5.34$ ;  $p = .024$ ;  $\eta^2 = .066$ ), without there being any significant differences between the 3 groups of self-centred participants (all  $ps > .172$ ).

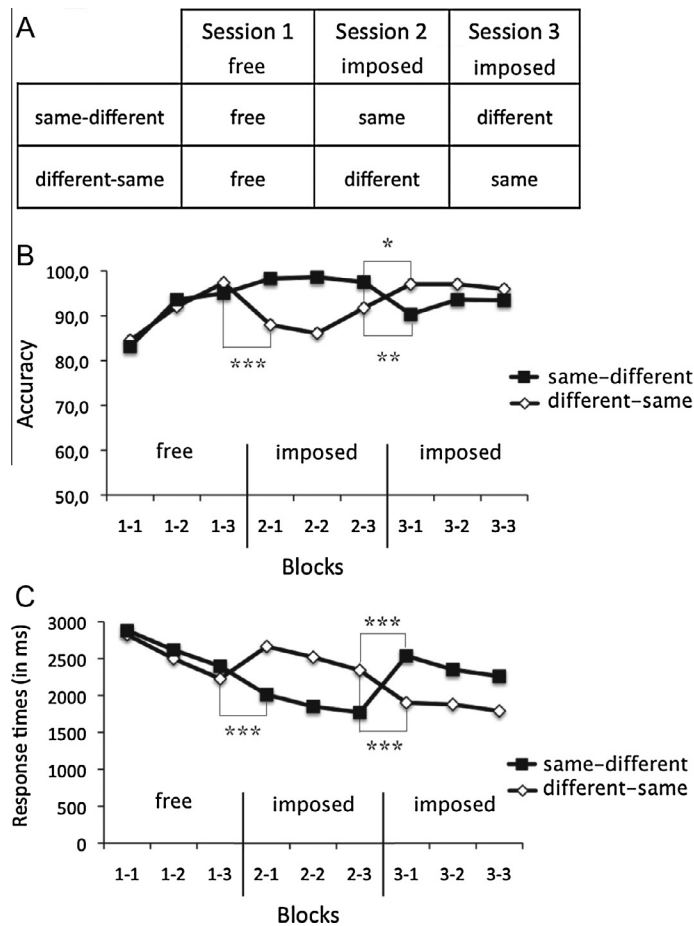
With respect to the RT data, the cost was significantly different from zero for all groups of participants (all  $ps < .001$ ). A one-way ANOVA showed that the cost was significantly smaller for the decentred than for the self-centred participants ( $F(1,76) = 6.75$ ;  $p = .011$ ;  $\eta^2 = .082$ ), without there being significant differences between self-centred participants (all  $ps > .116$ ).

### 3.4. Influence of visuospatial abilities

With respect to mental rotation, there was a significant negative correlation ( $r = -.34$ ;  $t(78) = 3.10$ ;  $p = .01$ ) between the score in the MRT and the cost of perspective change in accuracy: The greater the mental-rotation abilities, the smaller the cost. In addition, the trunk-centred (mean score = 26.7, SD = 7.9) and decentred participants (mean score = 28.5, SD = 7.3) showed greater mental-rotation abilities than the head-centred participants (mean score = 22.8, SD = 9.3;  $F(2,77) = 3.38$ ;  $p < .05$ ). With respect to visuo-spatial perspective taking, there was neither a significant correlation between the mean error in the OPTT and the cost of perspective change ( $r = -.07$ ;  $t(78) < 1$ ; *ns*) nor any significant differences between the groups ( $F(2,77) < 1$ ; *ns*).

## 4. Discussion

The study reported here was designed to evaluate whether perceivers have a natural perspective that constrains their perception of stimuli presented on their body surface. When perceivers are free to adopt any perspective that they want, 80% consistently adopt a self-centred perspective (50% trunk-centred, 30% head-centred) while 20% adopt a decentred one. The fact that



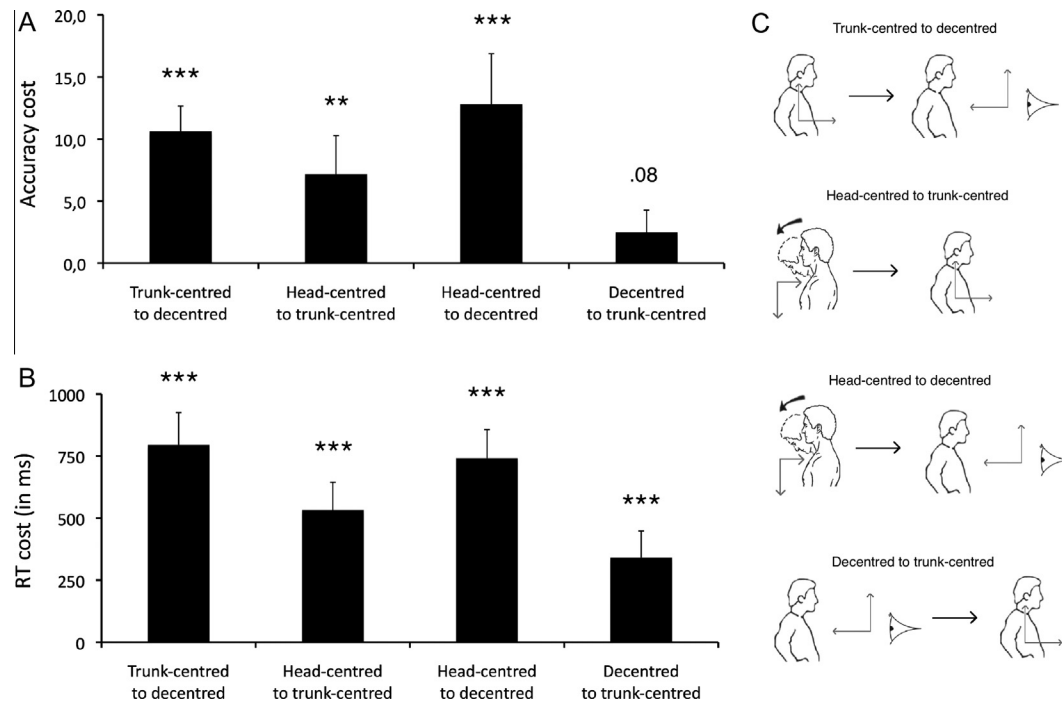
**Fig. 3.** (A) Illustration of the perspective imposed in each session for the two groups of participants. “Same-different” means that the imposed perspective was the same as in Session 1 for Session 2 and was different for Session 3. “Different-same” means that the imposed perspective was different for Session 2 and the same as in Session 1 for Session 3. Participants’ (A) accuracy (percentage correct) and (B) response times as a function of Block and Order of imposed perspectives (same-different, different-same). For both accuracy and response times, the performance decreased significantly when imposing a different perspective and increased significantly when returning to the same perspective as in Session 1. \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

recognition performance decreases with an unnatural perspective and returns to its previous high level with a natural one supports the view that the perspective adopted is natural. In addition, when a different perspective was imposed, the participants made more errors corresponding to the adoption of their natural perspective than other errors (see Quality of responses in the [Supplementary Materials](#)). Some of the observers perceive spatial relations better from a self-centred perspective whereas others perceive better from a decentred perspective. However, the greater cost for self-centred than decentred participants shows that decentred perceivers adopt more easily an unnatural perspective than self-centred perceivers.

The fact that the cost of adopting an unnatural perspective was correlated with mental rotation but not with visuo-spatial perspective taking also supports the existence of a natural perspective. When an unnatural perspective was imposed, the participants may have kept their natural perspective and then have mentally rotated the perceived letter rather than to really adopt the unnatural perspective (see [Surtees, Apperly, & Samson, 2013](#), for the role of mental rotation in perspective taking), possibly explaining the slowing of RTs. The possible involvement of mental rotation when an unnatural perspective is imposed raises the question of what adopting a spatial perspective really means. It could just involve specifying the spatial coordinates that result from the perspective that is being taken (i.e., the left–right orientation of someone else’s body space) without imagining the self being located at the decentred location.

In the present study, only 20% of the participants adopted a decentred perspective, which is less than reported in previous studies presenting ambiguous symbols on the stomach (50% in [Sekiyama, 1991](#); 71% in [Parsons & Shimojo, 1987](#)). This difference can be explained by the fact that the symbols were presented by means of a matrix of tactile vibrators, thus avoiding the major bias that may have been present in previous studies toward adopting the experimenter’s perspective (since they drew the stimuli on the participant’s skin). Nonetheless, even though the experimenter was not present in our experimental setup, some people appear to prefer adopting a decentred perspective, thus suggesting that this perspective may be their default perspective. In future work, it would be interesting to quantify such natural decentring by directly comparing the probability of decentring for tactile devices and for experimenter-drawn stimuli.

Self-centred perspectives can be centred on the stimulated surface (the trunk) or on a central body part (the head). [Sekiyama \(1991\)](#) has demonstrated that this head-centred perspective is adopted only when the corresponding bending-forward movement of the head toward the stimulated surface is possible. This head-centred perspective may thus involve a kind of visual strategy. The fact that some perceivers naturally adopt such a visual perspective may be explained by their lower spatial abilities. It would also be interesting to evaluate whether blind individuals adopt this head-centred perspective less frequently than do sighted individuals. Similar proportions of self-centred and decentred perspectives have been reported in blind and sighted people ([Shimojo, Sasaki,](#)



**Fig. 4.** Cost in (A) accuracy (percentage correct) and (B) response times when the different perspective was imposed on participants, as a function of the natural and the imposed perspectives. For example, ‘trunk-centred to decentred’ means that the cost was computed for the natural trunk-centred participants when adopting the decentred perspective. Error bars represent the standard errors of the means. The presence of asterisks indicates that the cost was significantly different from zero. \*\* $p < .01$ , \*\*\* $p < .001$ . (C) Schematic illustration of the natural and the imposed perspectives for the 4 different groups.

Parsons, & Torii, 1989) but the vertical inversion corresponding to the adoption of a head-centred perspective was not taken into account. Finally, the fact that vertical inversion was observed only when the head-centred perspective was adopted (see Quality of responses in the [Supplementary Materials](#)) reinforces previous observations that the top–bottom axis is less prone to confusion than the left–right axis (Oldfield & Phillips, 1983; Parsons & Shimojo, 1987; see also Farrell, 1979; Takano, 1998; Uehara, 2013). The vertical axis may thus be assigned before the horizontal axis when interpreting tactile symbols.

To conclude, the present study reveals that perceivers do not adopt the same perspective when interpreting ambiguous spatial information. Some perceivers prefer to perceive space from their own centred point of view whereas others prefer the point of view of another person. Even though decentred perspectives were adopted less frequently than self-centred ones, the natural adoption of a perspective decentred from the location of the body was observed (in 20% of the cases) in our study. Such spatial decentring also characterizes out-of-body experiences, where the self and the body are temporarily disconnected (Blanke, Landis, Spinelli, & Seeck, 2004; Ehrsson, 2007). Interestingly, a reinforcement of the processes anchoring the self to the body by galvanic stimulation of the vestibular system has recently been reported to bias participants toward self-centred perspectives in the ambiguous tactile symbol recognition task (Ferrè et al., 2014). The results of the present study, however, reveal that the adoption of a decentred perspective can reflect other processes than a sole distortion in the relation between the self and the body. Not only does it occur quite frequently, but it can also be considered as a stance crucial to know that other persons perceive the world in a differently way than we do and to understand how they perceive it.

### Acknowledgements

We thank Gaëtan Parseihan and Mathieu Dubois for their help in designing the apparatus. We also thank Eric Siéroff for giving

us the opportunity to test participants in his laboratory (Neuropsychologie du Vieillessement, EA 4468, Université Paris Descartes). This research was supported by the Agence Nationale de la Recherche (ANR-11-JSH2-003-1) and by the Mission pour l'Interdisciplinarité (Defisens, Supplé-Sens grant). GA is funded by grants from the Agence Nationale de la Recherche (ANR-11-JSH2-003-1) and from the Fondation des Aveugles de Guerre. CS is funded by the AHRC grant entitled Rethinking the Senses (AH/L007053/1). The Institut Jean Nicod laboratory finds support in ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL.

### Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.cognition.2015.12.006>.

### References

- Arnold, G., Spence, C., & Auvray, M. (2015). What perspectives do observers adopt when interpreting tactile stimuli? Reviewing the literature on the tactile ambiguous symbol recognition paradigm. *Psychonomic Bulletin and Review* (submitted for publication).
- Blanke, O., Landis, T., Spinelli, L., & Seeck, M. (2004). Out-of-body experience and autoscopia of neurological origin. *Brain*, *127*, 243–258.
- Duran, N. D., Dale, R., & Kreuz, R. J. (2011). Listeners invest in an assumed other's perspective despite cognitive cost. *Cognition*, *121*, 22–40.
- Ehrsson, H. H. (2007). The experimental induction of out-of-body experiences. *Science*, *317*, 1048.
- Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, *40*, 760–768.
- Farrell, W. S. (1979). Coding left and right. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 42–51.
- Ferrè, E. R., Lopez, C., & Haggard, P. (2014). Anchoring the self to the body: Vestibular contribution to the sense of self. *Psychological Science*, *25*, 2106–2108.
- Galinsky, A. D., Magee, J. C., Inesi, M. E., & Gruenfeld, D. H. (2006). Power and perspectives not taken. *Psychological Science*, *17*, 1068–1074.
- Harrar, V., & Harris, L. R. (2010). Touch used to guide action is partially coded in a visual reference frame. *Experimental Brain Research*, *203*, 615–620.



- Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, 32, 175–191.
- Ho, C., & Spence, C. (2007). Head orientation biases tactile localization. *Brain Research*, 1144, 136–141.
- Natsoulas, T. (1966). Locus and orientation of the perceiver (ego) under variable, constant, and no perspective instructions. *Journal of Personality and Social Psychology*, 3, 190–196.
- Natsoulas, T., & Dubanovski, R. A. (1964). Inferring the locus and orientation of the perceiver from responses to stimulation of the skin. *The American Journal of Psychology*, 77, 281–285.
- Oldfield, S. R., & Phillips, J. R. (1983). The spatial characteristics of tactile form perception. *Perception*, 12, 615–626.
- Parkinson, J., & Khurana, B. (2007). Temporal order of strokes primes letter recognition. *Quarterly Journal of Experimental Psychology*, 60, 1265–1274.
- Parsons, L. M., & Shimojo, S. (1987). Perceived spatial organization of cutaneous patterns on surfaces of the human body in various positions. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 488–504.
- Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition*, 47, 1–24.
- Sekiyama, K. (1991). Importance of head axes in perception of cutaneous patterns drawn on vertical body surfaces. *Perception & Psychophysics*, 49, 481–492.
- Shimojo, S., Sasaki, M., Parsons, L. M., & Torii, S. (1989). Mirror reversal by blind subjects in cutaneous perception and motor production of letters and numbers. *Perception & Psychophysics*, 45, 145–152.
- Surtees, A., Apperly, I., & Samson, D. (2013). Similarities and differences in visual and spatial perspective-taking processes. *Cognition*, 129, 426–438.
- Takano, Y. (1998). Why does a mirror image look left-right reversed? A hypothesis of multiple processes. *Psychonomic Bulletin and Review*, 5, 37–55.
- Thirioux, B., Jorland, G., Bret, M., Tramus, M. H., & Berthoz, A. (2009). Walking on a line: A motor paradigm using rotation and reflection symmetry to study mental body transformations. *Brain and Cognition*, 70, 191–200.
- Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition*, 110, 124–129.
- Uehara, I. (2013). Left-right and up-down mirror image confusion in 4-, 5- and 6-year-olds. *Psychology*, 4, 736–740.
- Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, 47, 599–604.
- Vogeley, K., & Fink, G. R. (2003). Neural correlates of first-person-perspective. *Trends in Cognitive Sciences*, 7, 38–42.
- World Medical Association (1991). Declaration of Helsinki. *Law, medicine & health care: A Publication of the American Society of Law & Medicine*, 19, 264–265.

# Cognition overrides orientation dependence in tactile viewpoint selection

Jessica Hartcher-O'Brien<sup>1,2</sup> · Malika Auvray<sup>1</sup>

Received: 14 December 2015 / Accepted: 8 February 2016  
© Springer-Verlag Berlin Heidelberg 2016

**Abstract** Humans are capable of extracting spatial information through their sense of touch: when someone strokes their hand, they can easily determine stroke direction without visual information. However, when it comes to the coordinate system used to assign the spatial relations to the stimulation, it remains poorly understood how the brain selects the appropriate system for passive touch. In the study reported here, we investigated whether hand orientation can determine coordinate assignment to ambiguous tactile patterns, whether observers can cognitively override any orientation-driven perspectives on touch, and whether the adaptation transfers across body surfaces. Our results demonstrated that the orientation of the hand in the vertical plane determines the perspective taken: an external perspective is adopted when the hand faces the observer and a gaze-centred perspective is selected when the hand faces away. Participants were then adapted to a mirror-reversed perspective through training, and the results revealed that this adapted perspective holds for the adapted surface and generalises to non-adapted surfaces, including across the body midline. These results reveal plasticity in perspective taking which relies on low-level postural cues (hand orientation) but also on higher-order somatosensory processing that can override the low-level cues.

**Electronic supplementary material** The online version of this article (doi:[10.1007/s00221-016-4596-6](https://doi.org/10.1007/s00221-016-4596-6)) contains supplementary material, which is available to authorized users.

✉ Jessica Hartcher-O'Brien  
jesshartcher@gmail.com; hartcher@isir.upmc.fr

<sup>1</sup> UPMC Univ. Paris 06, UMR 7222, ISIR, Sorbonne Universités, 75005 Paris, France

<sup>2</sup> Institut Jean Nicod, CNRS, EHESS, Ecole Normale Supérieure, 29 Rue d'Ulm, 75005 Paris, France

**Keywords** Touch · Viewpoint selection · Ambiguous pattern of stimulation · Coordinate assignment · Frame of reference

## Introduction

Spatial information about the objects in our environment is accessible not only through vision, but also through our senses of audition and touch (Vallbo and Johansson 1984). For instance, people can interpret the direction of stimulation when a visual or an auditory stimulus is displayed moving from left to right. This also holds true in touch when a body part is stroked from left to right. The question of how left and right are assigned, and therefore of how spatial information is taken from the environment, depends on the sensor that is used (Newell et al. 2001), with the mechanisms appearing to be more straightforward and established in vision than in touch.

In the case of visual viewpoint selection, when you watch your hand being stroked, the stroking pattern is initially coded in retinotopic coordinates. Depending on the task, either visual information is then either maintained in retinotopic coordinates for immediate action or it is transformed into external coordinates so that the observer can maintain a stable perceptual experience despite eye movements, changes in the position of the observer relative to the object, and object perturbations (Holway and Boring 1941; Wexler and Held 2005).

In the case of touch, the brain's viewpoint on stimuli is similarly not arbitrary; that is, when somebody strokes your hand, it is easy to distinguish whether they are stroking it from left to right or from right to left. However, unlike spatial processing in vision, at any given time there are potential contributions from multiple coordinate systems,

even when the task and posture do not change (Parsons and Shimojo 1987; Volcic et al. 2009). For example, determining the stroke direction across the surface of the hand can initially be done by coding the information in hand-centred coordinates, where left and right are assigned according to the spatial layout of the hand. However, the hand is not an independent system, but is always relative to the person's trunk, head, or eyes. In this case, the body-centred, cranio-topic, retinotopic, or gaze-centred (cranio- and retinotopic combined) coordinates can also be used to code and interpret the incoming tactile information. To summarise, concerning early stages of processing, visual information is initially coded in retinotopic coordinates and can then be transformed into external, cranio-topic, or motor coordinates depending on the task and goal of the observer, whereas for touch several coordinate systems already compete at early stages of somatosensory processing (Heed et al. 2015).

Theoretically, for an optimal processing of passive touch, i.e. when no explicit action is required, the most parsimonious interpretation of the stroking of the hand is to not remap the information into another coordinate system. For the brain to do so would be unnecessarily costly; thus, left and right would be assigned according to the hand's spatial layout. On the other hand, if one is actively engaged in the exploration of an object, then the incoming information may be most efficiently coded by transforming the incoming somatosensory information into external coordinates. By transforming the stimulation into external coordinates, the brain can maintain a stable representation of the object, despite distortions due to movement, and postural cues. However, the theoretical parsimonious approach of coding information in hand-centred coordinates is not necessarily what has been observed (Parsons and Shimojo 1987; Volcic et al. 2009). The question then becomes what are the factors that determine this coordinate assignment process for passive touch?

Given the ability of the hand to explore objects from different vantage points, independent of the posture of the body, this surface may well exhibit unique reference frame selection relative to other body surfaces. Several studies approached the question of what perspectives may be taken on somatosensory stimuli on the hand (Azañón and Soto-Faraco 2008; Holmes 2014; Prather and Sathian 2002; Volcic et al. 2009). In these studies, coordinate selection is based on tests that give changes in reaction time between the stimulus/hand placed in a baseline coronal position and when it changes to a different position (Prather and Sathian 2002). The increase in RT is an indication of the cost of changing reference frame (Volcic et al. 2009). Note that the studies that behaviourally assess reference frame selection in passive touch (without motor activity or movement) on other body parts also used indirect methods such as RTs (Farrell and Robertson 1998; Schicke and Röder 2006; Torralbo et al. 2006).

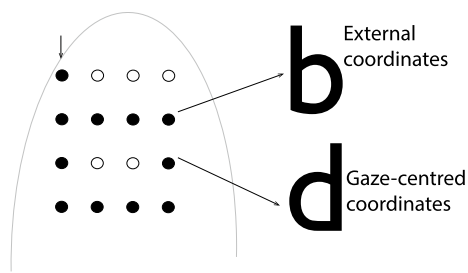
The exception, which provides a promising method to directly investigate viewpoint selection, was to use an ambiguous tactile symbol paradigm, which consists of grapheme stimuli (letters and numbers) drawn on the body surface. These graphemes potentially have several interpretations until the observer selects one coordinate system with which to interpret them. For instance, the letters 'b', 'd', 'p', and 'q' displayed on the skin are ambiguous until the observer takes a perspective and assigns coordinates to the cutaneous input, and consequently perceive one of the letters. Parsons and Shimojo (1987), for instance, used this paradigm and traced the graphemes on multiple body surfaces when the relative position and orientation of the surface was varied. When the experimenter traced the tactile patterns on the palm of the hand, stimuli were perceived relative to the position and orientation of the stimulated surface with respect to the body torso. Parson and Shimojo's results suggest the use of external frames of reference that are independent of the hand or head and that are associated with stimuli presented on the hand, that is, for Parsons and Shimojo, according to the experimenter-defined coordinates. However, in Parsons and Shimojo's study (and in other studies using this paradigm on other body surfaces that the hand, e.g. Corcoran 1977) the graphemes were always drawn manually by the experimenter. As a consequence, the transformation into experimenter-defined coordinates could be solely due to the influence of the experimenter himself biasing the participants' responses.

In our study, any experimenter-induced bias was removed from the equation by using a small Braille pin display device to present the directional cutaneous patterns (see Fig. 1). Our study was divided into three tasks. The first task investigated a similar question to Parsons and Shimojo's (1987) study: that is, whether the orientation of the hand could determine the mind's perspective on touch—however with our device and standardised procedure. However, in the case of our study, we were curious about interpretation of touch to the fingertip, not the palm (as investigated by Parsons and Shimojo 1987). We thus explored, for the first time, whether viewpoint selection could be overridden by adapting to the mirror-reversed perspective (in the second task). In a third task, we then measured whether this adapted perspective would generalise across non-adapted fingers and across the body midline.

## Methods

### Participants

Fourteen naïve participants (nine males) took part in the study. The average age was 26.8 years. The participants provided their written consent prior to the commencement



**Fig. 1** Cutaneous fingertip stimulation pattern and two potential interpretations depending on the reference frame selected. When the stimulation pattern *b* is presented on the device surface (start point and direction of stimulation indicated by the *arrow*), this pattern is interpreted as a *b* if the spatial coordinates are assigned according to the external object's coordinates. However, this example pattern is ambiguous around the vertical axis with a mirror-reversed perspective giving an interpretation of the pattern as the letter *d*. Here, we define the object-centred coordinates as 'external', i.e. according to the spatial layout of the device surface. Given the position of the device in front of the observer, the mirror reverse of this assignment is consistent with craniotopic as well as retinotopic coordinates. Because our task does not provide the means to disentangle the two potential coordinate systems, we have adopted the description used by [Harrar and Harris \(2009\)](#) where they define a merged coordinate system of head and eye as 'gaze-centred'

of the experiment and were given €8 for their participation. The experiment took approximately 1 h to complete and was performed in accordance with the ethical standards laid down in the 1991 Declaration of Helsinki.

### Apparatus and stimuli

Stimuli were produced using TACTOS ([Gapenne et al. 2003](#)), a custom-built device that generates tactile stimuli within a  $4 \times 4$  matrix design. The  $4 \times 4$  pin-matrix display measures 2 cm in width and 2.5 cm in height. When activated, each pin extends 2 mm from the display surface. Observers sat with their head approximately 60 cm from the device, which was Velcro-taped to a stable frame. During the experiment, observers wore an eye mask to stop visual information about the device influencing their judgments. Tactile signals were on/off activations of the Braille pins. The letter stimuli were formed by sequential activations of the pins as shown in [Fig. 1](#). The rate of activation was 20 mm/s. The stimulus lasted approximately 1000 ms in total with each pin activation corresponding to 100 ms. Each letter was defined according to the pattern of stimulation on the device surface.

### Procedure

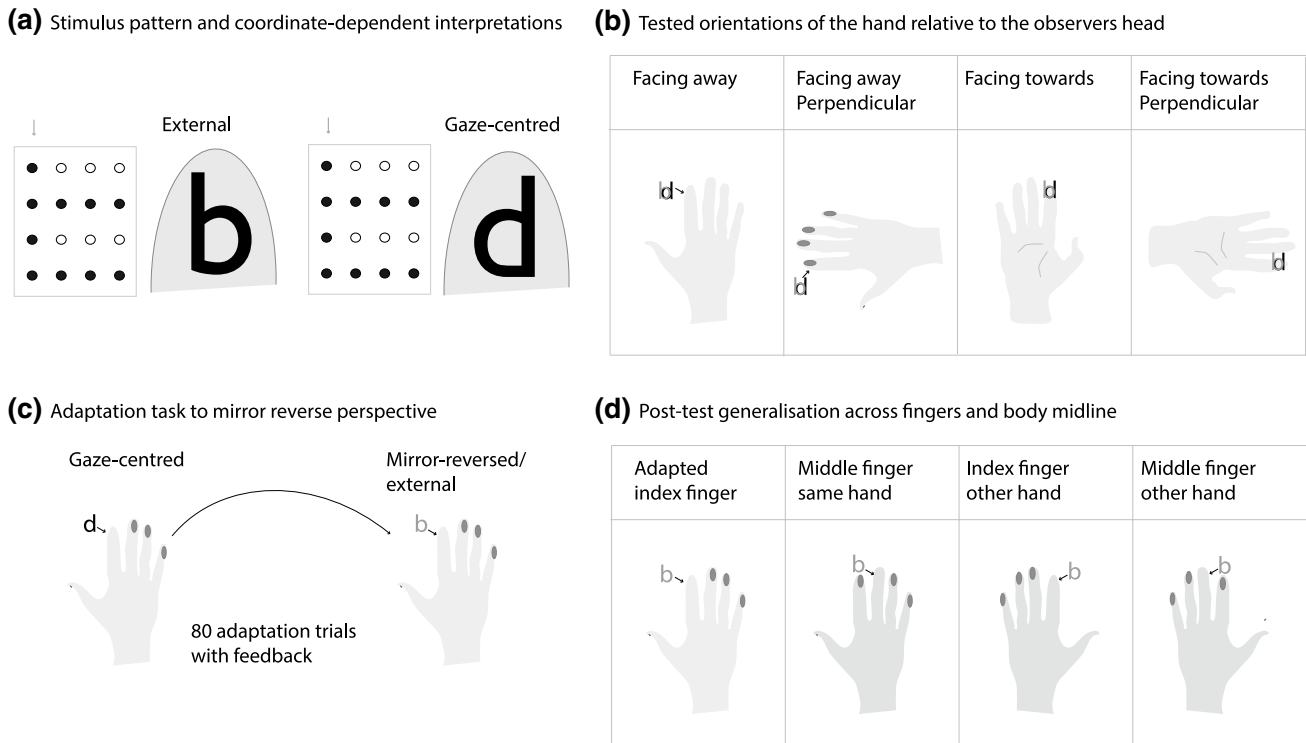
The experiment was divided into three tasks: first, an orientation manipulation, followed by a verification and an adaptation phase, and finally a post-test generalisation in

perspective taking task (see [Fig. 2d](#)). The first task was divided into four blocks, one for each of the four hand orientations. The participants sat and first fixated a central fixation point on the wall in front of them. They placed the volar surface of their index fingertip, in the orientation indicated by the experimenter, against the device display surface. The fingertip of the dominant hand of the observer was stimulated on the volar surface. The blocked, postural manipulation did not alter the surface stimulated but rather the orientation of the surface relative to the head of the observer and the front/back surface of the device. The participants were then blindfolded to ensure that visual information did not bias their responses. The device was attached to a stable frame so that the participants could press their fingertip against the active surface of the device. Each trial consisted of a single sequential tactile pattern, randomly selected from the stimulus list: 'b', 'd', 'p', or 'q'. Each letter began from the stem and consisted of a continuous motion ending with the body of the letter. We used a four-alternative forced-choice procedure in which the participants indicated which of the four letters they perceived, by pressing the corresponding key on the keyboard. The inter-trial interval varied randomly between 1000 and 2000 ms. The participants were able to respond during the stimulus presentation and up to 3000 ms after stimulus offset. They received no feedback as to the correctness of their responses. Each stimulus was repeated ten times in random order for each blocked hand orientation (i.e. 160 trials in total).

Task 2 consisted of an initial verification of the orientation dependent perspective and then an adaptation phase in which participants adopted orientation 1, as represented in [Fig. 2b](#), and were asked to reverse the assignment of left/right relative to their responses in the first experiment. That is, they were required to respond as though they were looking at the cutaneous pattern from the perspective of the device. Correct/incorrect feedback was provided over the 80 adaptation trials, with 20 repetitions of each of the four cutaneous letter patterns.

The third task consisted of four post-test blocks, one for each of the generalisation surfaces tested (see [Fig. 2d](#)). In this phase, the strength and generalizability of the adapted perspective was verified. In this post-test phase, the task was identical to that of the first part of the experiment. However, the hand orientation did not change across blocks: the hand was kept in the 'facing-away' orientation, but we changed the surface that was stimulated: the adapted index finger surface and three non-adapted finger surfaces (see [Fig. 2d](#)).

The dominant hand of the observer was used during the pretest and adaptation phases of the experiment. In the post-test phase, both hands were stimulated. The stimulated region was always the volar surface of the



**Fig. 2** Cutaneous patterns, potential perspectives, and fingertip orientations adopted during the experiment. **a** The device pin activation pattern for the letter *b*. The interpretation of the pattern as a *b* required assigning external coordinates to the activation pattern. An interpretation for a *d* would involve assigning the reverse, gaze-centred, coordinates. **b** The four surface orientations tested during task 1: orientation manipulation. Despite changes in the orientation of the

surface relative to the head and the device front/back surface, it is always the volar surfaces of the fingertip that were stimulated. The orientations are given with reference to the head of the observer. **c** Illustration of the adaptation phase used in task 2. **d** Fingers tested during the post-adaptation phase. Orientation of the fingertip was held constant

fingertip. During the first two phases of the experiment (i.e. the pretest and the adaptation phase), the index finger of the dominant hand was used. In the generalisation post-test phase, however, the volar surface of the two index fingers and middle fingers was stimulated, as shown in Fig. 2d.

**Results**

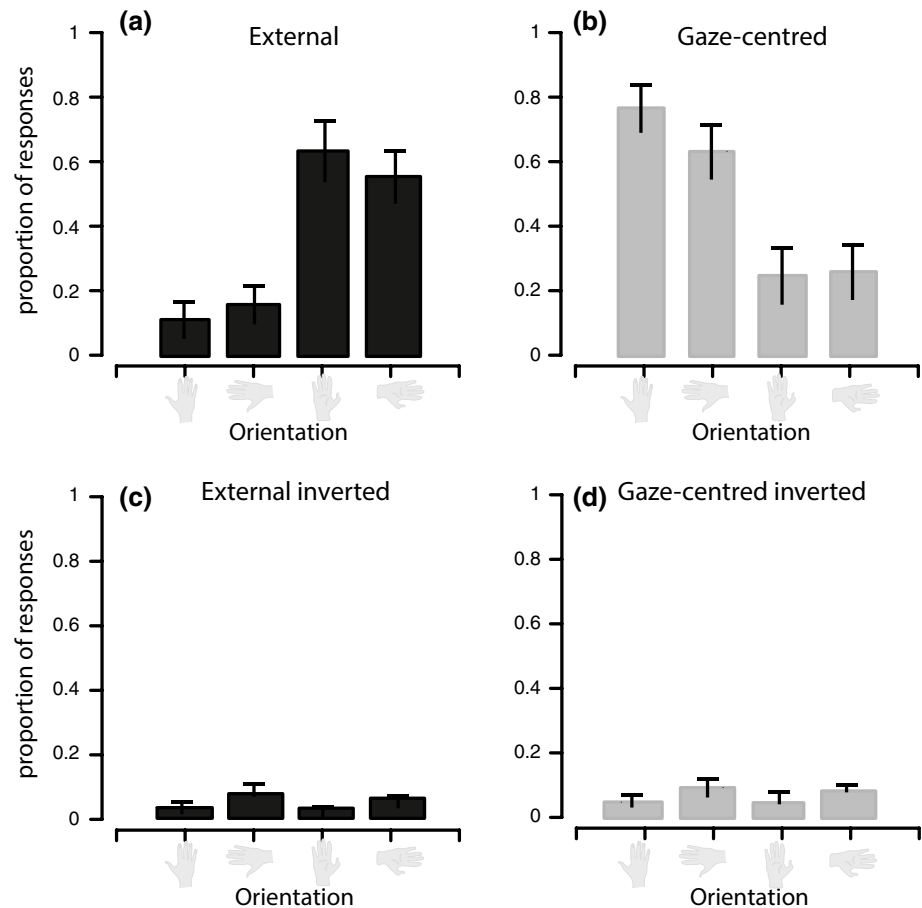
The participants' responses were transformed from letter identification estimates to proportions of external or gaze-centred (i.e. mirror-reversed) coordinates. External coordinates were defined as imposing spatial relations onto the pattern of stimulation according to the external device coordinates. In this case, the proportion of responses corresponding to this perspective was calculated as the number of 'b'/'d'/'p'/'q' responses consistent with a 'b'/'d'/'p'/'q' pattern on the device surface. Gaze-centred coordinates were defined as the mirror reverse of this assignment, with left/right, up/down being assigned according to the eyes/head of the participant. Thus, the response observed

indicates the perspective adopted by the participant during the tactile stimulation.

For the first task (orientation manipulation), we analysed the proportion of responses consistent with the participants adopting each of the four possible perspectives: external, gaze-centred, external but inverted, and gaze-centred but inverted. A one-way repeated-measures ANOVA was conducted on the different letter stimuli (four levels) which failed to reveal any significant difference in the perspective taken across the four letters [ $F(3,13) < 1, p > 0.05$ ]. Consequently, the data were averaged across the different cutaneous letter patterns.

A repeated-measures ANOVA, conducted on the participants' responses with the four-level orientation factor, was significant [ $F(3, 33) = 53.22, p < 0.001$ ]. Post hoc pairwise comparisons with Bonferroni's correction showed significant differences between orientations 1 and 3 (see Fig. 2b), [ $t(26) = 4.37, p = 0.001$ ] and between orientations 1 and 4 [ $t(26) = 4.34, p = 0.002$ ]. There was no difference between orientations 1 and 2 [ $t(26) = 1.15, p = 0.25$ ], nor between 3 and 4 [ $t(26) = 0.09, p = 0.92$ ]. These results demonstrate that the participants'

**Fig. 3** Proportion of responses for each orientation, averaged across the four cutaneous patterns *b*, *d*, *p*, and *q* consistent with assigning **a** external, **b** gaze-centred, **c** external but horizontally inverted, **d** gaze-centred but horizontally inverted coordinates to the cutaneous pattern. The proportion of responses sums to one across the four possible perspectives. The error bars represent the standard error of the mean across participants

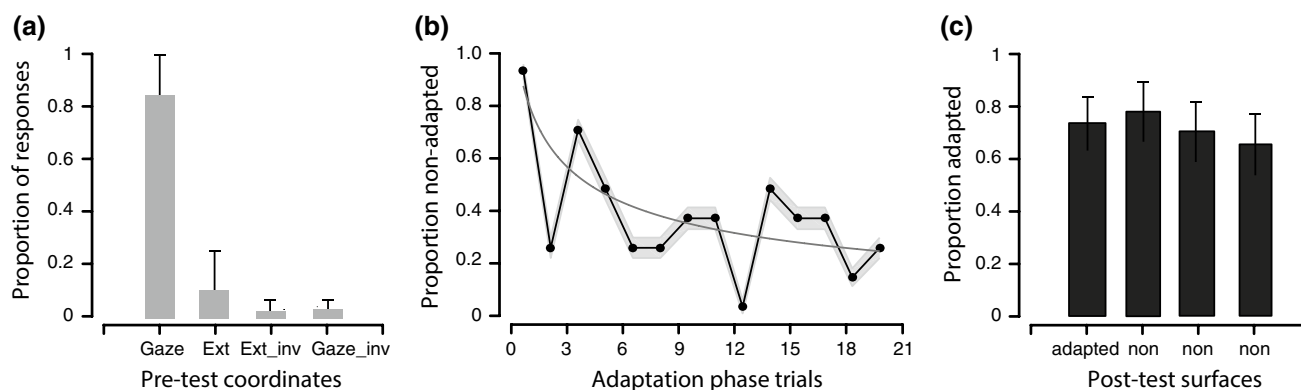


perspective depends on the orientation of the hand relative to the device with the differences being driven by the facing-away versus facing-towards positions, the upright versus perpendicular position having no influence on the results (Fig. 3a). In other words, the participants' perspective depends on the orientation of their hand along the vertical axis. With respect to the adopted perspective, when the hand is facing towards the participant, when both upright and perpendicular to the body midline, an external perspective is adopted. When the hand is oriented facing away from the participant, a gaze-centred perspective is taken on the tactile patterns of stimulation. The inverted perspectives were adopted significantly less than chance [ $t(13) > 1$ ,  $p < 0.012$ ], see Fig. 3c, d.

With respect to task 2, we analysed what happens during the pretest, the adaptation phase, and the post-test of the experiment, respectively. During the pretest, the proportion of responses was consistent with the gaze-centred perspective predicted from the results obtained during the pre-adaptation task (see Fig. 4a). The gaze-centred perspective was mainly adopted (82 % of the trials), and the three other perspectives occurred significantly less often than chance [ $t(11) = -2.49$ ,  $p = 0.95$ ,  $t(11) = -20.28$ ,  $p = 1$ ,  $t(11) = -16.69$ ,  $p = 0.9$ ], see Fig. 4a.

During the adaptation task, the perspective is hypothesised to shift from the initial perspective and converge towards the adapted perspective as a function of time (see Fig. 4b). Figure 4b represents the proportion of responses consistent with the non-adapted perspective across trials, averaged across the 14 participants and across the four letters. To understand the adaptation rate for novel perspective taking in touch, we fit a power function to the adaptation data. The function has a slope of 0.8489 (95 % CI 0.6912, 1.007) and an intercept of  $-0.409$  (95 % CI  $-0.5315$ ,  $-0.2864$ ), with  $R^2 = 0.7807$ ,  $p < 0.01$ . From the fit, we can predict that saturation in the amount of adaptation is expected after about 100 adaptation trials where the function reaches a local minimum. The change in response over time during adaptation demonstrates that our feedback (correct/incorrect response) modified the responses as to which letter was felt on the fingertip.

With respect to the generalisation post-test, a repeated-measures ANOVA was conducted on the participants' responses with one factor: finger stimulated (four levels). The analysis did not reveal significant differences in the perspective taken for any of the surfaces tested, the adapted one or the three non-adapted [ $F(3,33) = 1.938$ ,  $p = 0.14$ ]. Moreover, single-sample  $t$  tests, with Bonferroni's



**Fig. 4** Proportion of responses consistent with the natural and adapted perspectives, before, during, and after the adaptation task. **a** The proportion of responses consistent with the predicted, gaze-centred perspective, and the three other non-predicted ones, averaged across observers. The *error bars* represent the standard error of the mean. **b** The change in the response pattern during adaptation (averaged across observers and across the *four letters*). That is, adaptation phase trial 1 corresponds to four trials involving each of the *four letters b, d, p, q*). The variance (*shaded error bar*) represents

the standard error across observers. The polynomial fit to the proportion non-adapted responses, represented by the continuous *grey line*. **c** The proportion of responses in the post-test (averaged across *observers* and *letters*) consistent with the adapted perspective for the four surfaces: the adapted fingertip, the non-adapted middle finger of the adapted hand, the homologous finger of the non-adapted hand, and the middle finger of the non-adapted hand. The *error bars* represent the standard error of the mean

correction applied, revealed that the transfer effect is significantly higher than chance [ $t(11) > 2.5$ ,  $p < 0.002$ ] for all surfaces tested. Thus, the learnt perspective, consistent with mapping the stimulation into external coordinates, remains active for the adapted finger and generalises across the non-adapted, middle finger of the same hand and across the body midline to the homologous and middle finger of the non-adapted hand (Fig. 4c). These results show that the orientation cues can be overridden by training to the reverse perspective: causing participants to map the cutaneous information into external coordinates.

## Discussion

Our study explored whether postural cues can account for the mind's perspective on touch, whether the perspective is malleable and can be updated to a novel perspective via adaptation, and finally whether adaptation of one surface transfers to non-adapted surfaces.

## Orientation

The first main result to emerge from our study is that the perspective taken on ambiguous tactile patterns presented to the fingertip is dependent on the posture of the hand, i.e. its orientation. The results show a dependence on orientation around the vertical axis, and a rotation of  $90^\circ$  from the vertical to the horizontal axis did not influence the results.

Parsons and Shimojo (1987) previously explored whether the perspective taken on stimulation of the palm

and dorsal regions of the hand is affected by orientation of the hand. Observers in their study adopted seven different hand orientations. For orientations in which the hand was in front of them, the observers interpreted the cutaneous pattern primarily according to external coordinates for both surfaces of the hand. These results are consistent with interpreting the cutaneous pattern according to the experimenter's viewpoint. However, as was outlined in the introduction, in Parsons and Shimojo's experiment, the fact that the experimenter manually traced the letter on the participant's body surface might have biased their responses towards the experimenter's point of view. This would appear to be the case, given that in our study we have removed this potential bias by using an automatised presentation of the stimuli, and our results reveal that when the hand is in front of the observer, tactile information is transformed both into external coordinates and into gaze-centred coordinates.

With respect to the two coordinate systems that we observed, it should be noted that it is not unusual that touch on the hand is coded into coordinates other than hand-centred. Indeed, viewpoint selection in touch is a non-straightforward process because not only are there multiple different body surfaces, but the position and orientation of the surface relative to the rest of the body, the eyes, and the object itself strongly influences the reference frame adopted (Parsons and Shimojo 1987; Volcic et al. 2009). This leads to the interdependence of several different reference frames, namely those based on somatosensory/body, and external world coordinates (Harrar and Harris 2009). Several studies have suggested that tactile information on different surfaces can also be remapped into a gaze-centred,

or a motor coordinate system (Harrar and Harris 2009; Mancini and Haggard 2014). In addition to the multiple potential coordinate systems available for coding touch, we assumed that the stimulation on the fingertip may be influenced further by the fact that the fingers can move independently to the rest of the hand (functionally driven) and may therefore recruit additional coordinate systems for decoding stimulation, to that used for interpreting touch to the rest of the hand.

### Adaptation

The second result to emerge from our study concerns the plasticity in the perspective taken as a result of adaptation. When the participants held the orientation of their hand constant and were adapted to a mirror-reversed perspective, as compared with their initial choice for that orientation in task 1, the interpretation of the ambiguous cutaneous patterns was updated to the adapted (external) perspective. The results of this adaptation phase demonstrate that the mind's perspective on touch is malleable and plastic in the face of feedback. Note that the adaptation to the reverse perspective occurred rapidly, during the 80 trials of adaptation (20 for each cutaneous pattern). The fact that simple correct–incorrect feedback can override the low-level postural cue of orientation supports the idea that there is already an ambiguity in terms of which coordinate system might be applied to the pattern of stimulation on the skin, and given that the brain must resolve an existing conflict between multiple possible coordinate assignments for touch, a higher-order learning factor can therefore easily override low-level cues (Corcoran 1977; Parsons and Shimojo 1987; Volcic et al. 2009).

### Transfer of learning

The results from the third task of our study demonstrate that the adapted perspective transfers across non-adapted fingers and across the body midline (i.e. the other hand). This, in conjunction with the fact that 'correct'/'incorrect' feedback can mediate the adopted perspective (overriding orientation cues), suggests that the process we measured here is a higher-level perspective taking, rather than an automatic coordinate assignment process linked to the receptive fields of the stimulated surfaces (Harrar et al. 2014). The strength of the transfer did not decrease significantly when tested on non-adapted surfaces, which is interesting because it suggests that the adapted perspective relies on something more than perceptual factors (Spengler et al. 1997). Moreover, RT data (see supplementary material) suggest that the process we observe is not solely the result of mental rotation, given that the difference between RT data in the pre- and post-test phases is only

approximately 8 ms. While significant, this is a different order of magnitude to that observed in mental rotation tasks (e.g. Just and Carpenter 1985).

For learning of tactile patterns on different body surfaces, i.e. the trunk and leg of the observer, Arnold and Auvray (2014) have also found a transfer of learning to non-adapted surfaces. The authors suggest that their effect is driven by the organisation of the somatosensory cortex. However, in our case where the transfer is about reference frame rather than patterns of stimuli, the fact that the learning transfers across the body midline suggests that the transfer at stake occurs at later stages in the processing pathway, given the lateralisation of early somatosensory processing (Boven et al. 2005). The ability to train a person to adopt the mirror-reversed perspective may be due to the fact that the stimuli themselves are high level and therefore recruit higher-order processing regions.

The transfer results have implications for training with novel haptic interfaces, suggesting that it is possible to remap, with a limited amount of training, the local cutaneous pattern of stimulation into different coordinates for the purposes of experiencing a three-dimensional world and a novel perspective from your own gaze-/body-centred point of view.

### Implications of the results

For the results obtained across the three tasks of our study: the perspective on touch adopted as a function of hand orientation, adaptation, and transfer, we could speculate that a similar level of processing is actually involved, given that the cutaneous patterns were not found to be coded in hand-centred coordinates for any orientations. That is, our results can be explained by higher-order processes probably involving brain areas not organised topographically (see also Harrar et al. 2014). It is clear that the observer's interpretation of patterns of stimulation on the skin is influenced by a number of factors, including competing coordinates systems (Parsons and Shimojo 1987), object identity (Arnold and Auvray 2014), the goal of the observer (Oldfield and Phillips 1983), and the observer's natural preferences (Arnold et al. 2016). What we observe with passive stimulation of tactile letters is that the competition between these different factors is resolved by stable high-level factors. We can quantify the perspective at the fingertip as being driven by postural cues but informed ultimately by learning, mental rotation, and feedback.

This can be accounted for by the fact that in order to maintain a stable experience of our environment, despite our ability to move our sensors relative to the environment, we need to be able to take into account the change in our own sensor positions (Hartcher-O'Brien and Auvray 2014). Thus, the orientation dependence observed in the first task of the current



study highlights one way in which the brain tries to account for changes in sensor position, i.e. to map information into gaze-centred and external coordinates. For visual processing, the visual system remaps information into external coordinates in order to maintain a stable representation across eye and stimulus movements (Epstein 1977). Here, our study shows how the brain updates its perspective on cutaneous stimulation during changes in hand orientation, transforming the patterns into non-hand-based coordinates. That is, in order to maintain a stable representation of the environment, the brain selects predominantly gaze-centred and external coordinates during the processing of passive tactile information.

To conclude, the coordinate system chosen to interpret tactile objects is a complex interplay between the object position in space, on the skin, and the orientation of the body part receiving the stimulation relative to other body surfaces (Parsons and Shimojo 1987). In our study, in no case was the hand-centred perspective adopted over the external and gaze-centred perspectives. The importance of hand orientation in perspective taking in everyday life, as demonstrated in the first part of the experiment, is highlighted in the example of how your brain uses the orientation of your hand to infer the surface (top or bottom) of a table: when you rest your hand on the top of a table, your perspective on the table's surface is unambiguous and primarily coded from the orientation of your hand. You know that your hand is in contact with the tabletop. If you rotate your hand 180° and establish contact with the table again, it is clear that you are in contact with the lower surface of the table. In any construction of external objects and space, it is important to provide coordinates that can be used to determine the spatial relations between objects and our senses. When we touch objects or when we are touched, both physical and physiological constraints influence our interpretation of the objects we perceive. Nevertheless, cognitive factors can override physiological constraints as seen in the adaptation of the mind's perspective on touch.

**Acknowledgments** JHO was funded by the Fyssen Foundation. MA was funded by a grant from the Agence Nationale de la Recherche (ANR-11-JSH2-003-1). The Institut Jean Nicod laboratory receives financial support via the ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL.

## References

- Arnold G, Auvray M (2014) Perceptual learning: tactile letter recognition transfers across body surfaces. *Multisens Res* 27(1):71–90. doi:[10.1163/22134808-00002443](https://doi.org/10.1163/22134808-00002443)
- Arnold G, Spence C, Auvray M (2016) Taking someone else's spatial perspective: natural stance or effortful decentring? *Cognition* 148:27–33. doi:[10.1016/j.cognition.2015.12.006](https://doi.org/10.1016/j.cognition.2015.12.006)
- Azañón E, Soto-Faraco S (2008) Changing reference frames during the encoding of tactile events. *Curr Biol* 18(14):1044–1049
- Boven RWV, Ingeholm JE, Beauchamp MS, Bickle PC, Ungerleider LG (2005) Tactile form and location processing in the human brain. *Proc Natl Acad Sci USA* 102(35):12601–12605. doi:[10.1073/pnas.0505907102](https://doi.org/10.1073/pnas.0505907102)
- Corcoran DWJ (1977) The phenomena of the disembodied eye or is it a matter of personal geography? *Perception* 6(3):247–253. doi:[10.1068/p060247](https://doi.org/10.1068/p060247)
- Epstein W (1977) *Stability and constancy in visual perception: Mechanisms and processes*. Wiley, New York
- Farrell MJ, Robertson IH (1998) Mental rotation and automatic updating of body-centered spatial relationships. *J Exp Psychol Learn Mem Cogn* 24(1):227–233. doi:[10.1037/0278-7393.24.1.227](https://doi.org/10.1037/0278-7393.24.1.227)
- Gapenne O, Rovira K, Ali Ammar A, Lenay C (2003) Tactos: special computer interface for the reading and writing of 2D forms in blind people. *Univers Access HCI Incl Des Inf Soc* 10:1270–1274
- Harrar V, Harris L (2009) Eye position affects the perceived location of touch. *Exp Brain Res* 198(2–3):403–410. doi:[10.1007/s00221-009-1884-4](https://doi.org/10.1007/s00221-009-1884-4)
- Harrar V, Spence CJ, Makin TR (2014) Topographic generalization of tactile perceptual learning. *J Exp Psychol Hum Percept Perform* 40(1):15–23. doi:[10.1037/a0033200](https://doi.org/10.1037/a0033200)
- Hartcher-O'Brien J, Auvray M (2014) The process of distal attribution illuminated through studies of sensory substitution. *Multisens Res* 27(5–6):421–441
- Heed T, Buchholz VN, Engel AK, Röder B (2015) Tactile remapping: from coordinate transformation to integration in sensorimotor processing. *Trends Cogn Sci* 19(5):251–258. doi:[10.1016/j.tics.2015.03.001](https://doi.org/10.1016/j.tics.2015.03.001)
- Holmes NP (2014) Hand-centred space, hand-centred attention, and the control. *Cortex* 25:189–220
- Holway AH, Boring EG (1941) Determinants of apparent visual size with distance variant. *Am J Psychol* 54(1):21–37
- Just MA, Carpenter PA (1985) Cognitive coordinate systems: accounts of mental rotation and individual differences in spatial ability. *Psychol Rev* 92(2):137–172. doi:[10.1037/0033-295X.92.2.137](https://doi.org/10.1037/0033-295X.92.2.137)
- Mancini F, Haggard P (2014) Perception: a motion after-effect for voluntary actions. *Curr Biol* 24(2):R70–R72
- Newell FN, Ernst MO, Tjan BS, Bühlhoff HH (2001) Viewpoint dependence in visual and haptic object recognition. *Psychol Sci* 12(1):37–42
- Oldfield SR, Phillips JR (1983) The spatial characteristics of tactile form perception. *Perception* 12(5):615–626. doi:[10.1068/p120615](https://doi.org/10.1068/p120615)
- Parsons L, Shimojo S (1987) Perceived spatial organization of cutaneous patterns on surfaces of the human body in various positions. *J Exp Psychol Hum Percept Perform* 13:488–504
- Prather SC, Sathian K (2002) Mental rotation of tactile stimuli. *Cogn Brain Res* 14(1):91–98
- Schicke T, Röder B (2006) Spatial remapping of touch: confusion of perceived stimulus order across hand and foot. *Proc Natl Acad Sci* 103(31):11808–11813. doi:[10.1073/pnas.0601486103](https://doi.org/10.1073/pnas.0601486103)
- Spengler F, Roberts TPL, Poeppe D, Byl N, Wang X, Rowley HA, Merzenich MM (1997) Learning transfer and neuronal plasticity in humans trained in tactile discrimination. *Neurosci Lett* 232(3):151–154. doi:[10.1016/S0304-3940\(97\)00602-2](https://doi.org/10.1016/S0304-3940(97)00602-2)
- Torralbo A, Santiago J, Lupiáñez J (2006) Flexible conceptual projection of time onto spatial frames of reference. *Cogn Sci* 30(4):745–757. doi:[10.1207/s15516709cog0000\\_67](https://doi.org/10.1207/s15516709cog0000_67)
- Vallbo AB, Johansson RS (1984) Properties of cutaneous mechanoreceptors in the human hand related to touch sensation. *Hum Neurobiol* 3(1):3–14
- Volcic R, Wijntjes MWA, Kappers AML (2009) Haptic mental rotation revisited: multiple reference frame dependence. *Acta Psychol* 130(3):251–259
- Wexler M, Held RM (2005) Anticipating the three-dimensional consequences of eye movements. *Proc Natl Acad Sci USA* 102(4):1246–1251

# The Process of Distal Attribution Illuminated Through Studies of Sensory Substitution

Jess Hartcher-O'Brien<sup>1,2,\*</sup> and Malika Auvray<sup>1,2</sup>

<sup>1</sup> Institut Jean Nicod, CNRS UMR 8129, Département d'Etudes Cognitives, Institut d'Etude de la Cognition, Ecole Normale Supérieure, Paris, France

<sup>2</sup> LIMSI CNRS, B.P. 133, Orsay-Cedex, 91400 France

Received 31 December 2013; accepted 30 June 2014

---

## Abstract

When we interact with objects in our environment, as a general rule we are not aware of the proximal stimulation they provide, but we directly experience the external object. This process of assigning an external cause is known as distal attribution. It is extremely difficult to measure how distal attribution emerges because it arises so early in life and appears to be automatic. Sensory substitution systems give us the possibility to measure the process as it occurs online. With these devices, objects in our environment produce novel proximal stimulation patterns and individuals have to establish the link between the proximal stimulation and the distal object. This review disentangles the contributing factors that allow the nervous system to assign a distal cause, thereby creating the experience of an external world. In particular, it highlights the role of the assumption of a stable world, the role of movement, and finally that of calibration. From the existing sensory substitution literature it appears that distal attribution breaks down when one of these principles is violated and as such the review provides an important piece to the puzzle of distal attribution.

## Keywords

Sensory substitution, distal attribution, prior of a stable world, movement, calibration, externalisation

## 1. Distal Attribution

As we move through the world, we are directly aware of external objects not the proximal stimulation they provide at our receptors (Crick and Koch, 1995; O'Regan, 1992). For instance, changes in air pressure or vibrations arriving at our cochlea are perceived as arising from an object in the environment (such as a moving car) rather than from the cochlea itself. In this instance, proxi-

---

\* To whom correspondence should be addressed. E-mail: jesshartcher@gmail.com

mal patterns of stimulation (e.g., vibrations in the cochlea) are perceived as having an external cause (e.g., the car). What are the cues necessary for this assignment to emerge?

### *1.1. Definition*

Distal attribution, also known as externalisation, occurs when we experience a pattern of stimulation and assign the cause of this stimulation to an object, stimulus, or event located in external three-dimensional space. This process allows us to make sense of the world around us. When proximal stimulation (i.e., stimulation received at the receptor surface) is not assigned an external cause, it gives rise to a mere sensation, not the perception of an object located at a distance. Therefore understanding how distal attribution emerges is key to understanding the distinction between sensation and perception.

### *1.2. The Problems in Investigating Distal Attribution*

The process of distal attribution is so automatic and present so early in life that it is extremely difficult to investigate it objectively or even to understand how it emerges (Bach-y-Rita, 2002; Loomis, 1992). Studies attempting to understand the phenomenon of distal attribution have therefore relied upon instances where distal objects give rise to novel proximal stimulation patterns, as happens with the use of sensory substitution devices (SSDs from now on) (Auvray *et al.*, 2005; Bach-y-Rita, 1972; Bach-y-Rita *et al.*, 1969; Epstein *et al.*, 1986; Loomis, 1992; von Békésy, 1955; White *et al.*, 1970).

To introduce them, SSDs were initially designed to assist or replace specific functions of a deficient sensory modality by providing corresponding information to an alternative sensory modality. This can be achieved by converting the stimuli normally sensed through the deficient modality (e.g., light for SSDs compensating for vision) into stimuli accessible to another sensory modality (e.g., tactile vibrations or sounds). Since their inception in the sixties various kinds of devices have been developed, tested, and shown to allow their users to behave to some degree as if they possessed the substituted sensory organ (see Auvray and Myin, 2009; Bubic *et al.*, 2010; Deroy and Auvray, 2012, for reviews). For instance, thanks to visual-to-auditory and visual-to-tactile conversion systems, blind individuals report being able to localize and recognize objects in three-dimensional space (e.g., Auvray *et al.*, 2007; Bach-y-Rita *et al.*, 1969; Levy-Tzedek *et al.*, 2012). Sensory substitution offers a novel way to investigate how distal attribution emerges because with sensory substitution devices the boundary between proximal- and distal-attribution is blurred. Such devices provide the unusual circumstance in which the perceiver initially experiences a proximal pattern of stimulation and, after training, perceives this same pattern as a distal event (Bach-y-Rita, 2002).

### *1.3. Review Objectives*

With SSDs, it is possible to observe how users transform the proximal stimulation rendered by the device into an experience of external objects. In such mediated perception (similar to the non-mediated kind), if we fail to assign a distal cause, we would no longer have a perceptual space corresponding to physical space, but we would be left with the sensation only. Therefore, as was mentioned above, understanding how distal attribution emerges, in both instances, is crucial to understanding perception itself. When distal attribution fails to emerge with SSDs, the users do not establish the corresponding hypothesis of a distal cause of the sensation (e.g., Bach-y-Rita, 1972; Epstein *et al.*, 1986). The literature to date has not directly addressed the link between prior beliefs of an external stable world and the process of distal attribution. Here we will explore the role of such assumptions in the emergence of distal attribution using SSDs. Furthermore we will identify three components essential for distal attribution to emerge: movement with its consequent feedback about the world, calibration of space, and the prior of a stable world. In turn this will allow us to reconcile how the use of distance perception as synonymous with distal attribution has changed our approach to testing distal attribution with SSDs. We would suggest that distance perception is a bi-product of the distal attribution process and that both phenomena depend on a prior of a stable world (e.g., Glennerster *et al.*, 2006; Knill, 2007).

## **2. Establishing External Space**

### *2.1. Problems of Establishing External Space with SSDs*

As was noted above, SSDs allow us to sense distal objects by providing novel proximal stimulation that corresponds to the external object at an alternative receptor surface. As such, the novel perceptual space that emerges is not necessarily a function of the same assumptions, regularities or spatial maps as those underpinning non-mediated perception. Through the SSD interface, observers have to re-learn the relationships between the experienced pattern of stimulation and the properties of the external objects. When distal attribution of proximal stimulation does not occur in SSD space, it is possible to see how prior beliefs fail to transfer and how the process of distal attribution breaks down. For example, if a proximal pattern of stimulation is experienced on the skin and is derived from interacting with a teacup 0.5 meters away, then the observer needs to re-establish the relationship that links the received proximal patterns to the teacup. If only the proximal pattern is experienced, i.e., without its link to the external object, then the perceiver does not experience external space and as a consequence no distance estimate can be derived.

On the other hand, if the observer has no reason to assume that the experienced pattern represents an object 0.5 meters away, then the mere experience of tactile vibrations, i.e., the lack of distal attribution, is the most parsimonious interpretation of the incoming stimulation. In a recent study, Siegle and Warren (2010) investigated the perceived distance of objects experienced through a minimalistic SSD. Participants were divided into two groups and performed the task under two different conditions: observers were asked either to focus their attention on the proximal stimulation on the body surface or, in the second condition, they were asked to focus their attention on the distal source. Blindfolded participants observed a target light using a device consisting of a single finger-mounted photodiode that drives tactile vibration on the back whenever the finger points at the target. After active exploration of their environment, participants had to move a reference object (experienced without the SSD) to match the perceived egocentric distance of the target. The participant group who were instructed to focus their attention on the distal object had higher performance than the group who focused their attention on the proximal stimulation during the task. However, while Siegle and Warren's study clearly showed a difference in performance between the two groups, this difference could be the result of attention rather than distal attribution per se. In non-mediated perception both externalisation and distance estimation depend upon the prior of an external world. Although the two can be conflated, as in Siegle and Warren's study, it is important to understand their relative contributions to our construction of an external space. In the following sections we will discuss what is known about the emergence of the concept of space in non-mediated perception and then turn to the role of such processes in SSD space.

### *2.1.1. The Emergence of Space During Non-Mediated Perceptual Processing*

We do not have direct access to the physical world. When we move our nervous system is provided with feedback about its current environment. Thus what we have access to are the relative changes in position of our sensory organs (e.g., our eyes or hand), and the consequent changes in sensory input. The consequence of movement is that our senses acquire a sequence of stimulation patterns. The information from this sequence is sufficient to determine the structure of 3D space (e.g., Glennerster *et al.*, 2006; Graham, 1989). Therefore we can experience space through the simple relationship between, for example, the distance the eyes have moved and concurrent changes in sensory stimulation. Thus our experience of an external world depends upon our ability to account for the effects of self-generated movement on changes in proximal stimulation. When the two are correlated in time, the sensory input can be experienced as objects in external space. In a Bayesian sense, the nervous system has a remarkable ability to detect statistical correlations between

sensory inputs and receptor states and it can use this information to form perceptual representations of the external world.

To minimise the cost of keeping track of the changes, the nervous system can establish the likelihood that a given sensorimotor relationship indicates a particular physical object and thus predict the influence of other exploration patterns. Previous research suggests that the cost of keeping track of the changes can also be minimised by integrating prior knowledge we have about the world and the physical laws that constrain objects therein (e.g., Glennerster *et al.*, 2006; Hayward, 2011; Loomis, 1992). To map proximal stimulation to objects in external space, the nervous system has to have an assumption that there is an external world. Inherent in this is the belief that the world is stable and that any physical object can only change its position (according physical laws) or its state (according to the laws of nature).

When space emerges from our interactions with the environment, the number of spatial dimensions that we perceive depends upon the sensor we use to detect the incoming information: perception of objects in depth is most easily accessed through vision, audition and haptics/kinaesthetic. In his seminal work, “The World of Touch”, Katz (1925) suggested that touch has many of the capabilities of a distal or ‘far’ sense, such as vision and audition (see also Krueger, 1970 for a summary of Katz’s work). However, spatial properties of objects, such as their distance from the observer, are readily available through the ‘far senses’ but not to the other senses and in this way our experience of space depend on the sensory receptor stimulated. For tactile perception, attribution to a distal source or to the body surface can differ depending on the circumstances: Active exploration via touch tends to promote distal attribution, whereas passively received tactile input to the skin are more likely to be experienced as coming from the body itself (Gibson, 1962; Katz, 1925). If visual information is converted into a tactile code then the ability to passively experience depth depends on remapping this property into a property accessible to the skin. Therefore, it is important to consider how the perception of external objects changes when we use a SSD. One issue is that in non-mediated perception, experience consolidates the mapping between the proximal stimulation and its distal object over time. In SSD space novel mappings emerge and need to be consolidated over time. Moreover, the information experienced via the device needs to be aligned with information coming from other sensory modalities.

### *2.1.2. Maintaining Spatial Relations when Going from Physical to Perceptual Space*

In order to survive, organisms need to correctly locate objects in space, orient to them, or flee from them. To aid in this response optimisation, humans and other animals have evolved spatial maps or fields that organise the incoming

information according to the spatial relations among physical objects. Each sensory modality has its own map and set of coordinates. The nervous system establishes spatial maps at low-level processing sites and all perceptual information is segregated according to these maps throughout the processing hierarchy. To be beneficial to behaviour, information from the different maps needs to be aligned such that a coherent object localisation emerges despite different sensory co-ordinates. Animal research indicates that neurons at multisensory sites such as the SC and the VPM cortex code the object location relative to the receptor used and furthermore that different sensory maps are aligned (e.g., Graziano *et al.*, 1999; Lee and Groh, 2009). When an auditory signal is moved in depth, neurons in the monkey's ventral premotor cortex are selectively tuned to respond maximally to specific locations in physical space (Graziano *et al.*, 1999). Through the selective responses a neural spatial map is established in the premotor cortex. Critically, these neurons, unlike those in A1 or visual areas, also respond to visual stimuli located in the same external co-ordinates. Animal and human research alike suggests that these maps are critical to our perceptual representation of space and spatial relations (Graziano *et al.*, 1999; King *et al.*, 2001; Moore and King, 1999; Poggio and Fisher, 1977).

We acquire information about the world from all of our senses. Our ears detect air pressure changes, our eyes monitor light patterns, our skin reports the pressure of objects impinging on it, and our nose and mouth measure the chemical content of our current environment. Despite the variety of routes that sensory information can take to get to our brains, we tend to experience the world as a unified entity. We are also capable of recruiting different types of sensory information for the same purpose, irrespective of which sensor was initially responsible for detecting the information. Given this, proximal stimulation through a SSD should potentially correspond to the experience of an object located in external space. Although distal attribution through SSDs has been the subject of many studies, most research has focused on the behavioural changes and therefore there is no information relating spatial map organisation to perception of spatial properties of the world during SSD use. It remains to be seen which maps will be recruited for coding the incoming information and how the interactions with the objects in 3D space are altered. One possibility is that perception mediated by a SSD, produces similar neural changes as those occurring in sensory deprivation (see Knudsen, 1999) and the consequent recruitment of a different sensory pathway to process incoming information. That is, the maps recruited to organise the proximal stimulation would correspond to those of the functional modality. In any case, for distal attribution to emerge the nervous system needs to adapt to the novel sensorimotor mappings and thus re-establish external space. Whether and how this changes the neural processing is not yet completely understood. The interesting aspect for the

current review is how the possible changes in sensory information would be organised to maintain spatial relations.

### **3. The Process of Distal Attribution**

The distinction between proximal stimulation and attribution of this stimulation to an external object requires the assumption of an external world (e.g., Epstein *et al.*, 1986). This may seem an obvious statement. Indeed it is. However, not addressing such assumptions alters how we define the key principles or mechanisms underlying distal attribution. Many studies have explored the question of how distal attribution arises during SSD-use. Early studies were based on qualitative reports (e.g., Bach-y-Rita *et al.*, 1969) whereas later studies used different measurement techniques, differing not only in their objectivity, but also in whether participants had direct knowledge of external space or only the sensory substitution space (e.g., Epstein *et al.*, 1986; White *et al.*, 1970). Each measurement technique has advantages and disadvantages. But here we focus on how embedding the assumption of an external world in the task itself, as opposed to defining space as purely an internal entity, alters the attribution process. We will address how valid the SSD data are in informing us about the process of distal attribution itself. We will then define the essential components influencing distal attribution. The factors involved in this process are: (1) the prior of a stable world; (2) the link between self-generated movement and the resulting sensory stimulation patterns and the additional laws of movement necessary for spatialisation (i.e., reversibility and interposition); and (3) the role of calibration in establishing and maintaining an experience of external space.

#### *3.1. The Problematic Aspects of SSD Measures for Understanding the Process of Distal Attribution*

There are numerous verbal reports from participants in SSD studies that detail the emergence of distal attribution after prolonged SSD-mediated experience (Bach-y-Rita *et al.*, 1969). That is, SSD-users report experiencing objects in the environment and this tends to dominate the proximal stimulation afforded by the device (Bach-y-Rita, 1972; Lenay *et al.*, 2003; Segond *et al.*, 2005). More particularly, SSD-users have indicated that, as a consequence of training, they no longer feel the tactile stimuli on their skin, where they occur, but interpret them directly as an object located in external space (see Bach-y-Rita, 2002). One well-cited verbal report comes from Guarniero, a blind philosophy student, after training with the TVSS: "... very soon after I had learned how to scan, the sensations no longer felt as if they were on my back, and I became less and less aware that vibrating pins were making contact with my skin. By this time objects had come to have a top and a bottom; a right side and a left;



but no depth — they existed in an ordered two dimensional space.” (Guarniero, 1974, p. 104).

One issue with verbal reports is that they cannot be verified in the way an objective measure can (e.g., a left–right judgment). Although these verbal reports indicated that distal attribution emerges in SSD space after training, there was no additional measure to specify how it emerged. That is, it remains unclear whether these reports reflect a genuine perceptual experience of distal objects or a cognitive inference based on perceived patterns in the proximal stimulation. The techniques used to address the question of distal attribution via SSDs have improved since the inception of the devices. In the following sections we will highlight the assumptions and cues to distal attribution that have emerged from the existing studies.

### *3.2. The Role of the Prior of a Stable External World*

To externalize proximal stimulation we need to assume the existence of an external world. Without this prior belief (or assumption) that the external world is a stable entity, neither distal attribution nor the explicit distance estimates of the distal object would be possible. As mentioned above, when we explore the world, we do not have direct access to the objects in space. What we have access to are the proximal patterns of activation and our own movements (Merleau-Ponty, 1945). The nervous system compensates for the distortions in incoming stimulation, due to movement, and manages to maintain the experience of a stable world and of the objects therein. The prior of an external, stable world is essential to this process. This prior posits that objects are stable. Inherent in this is that if an object changes it can only change its state (according to the laws of nature) or its position (according to the laws of physics). When a change is inconsistent with either positional or state changes in the object itself, the parsimonious assumption is that observer-generated movement must have caused the change in incoming information. These kind of changes can easily be correlated with the efference copy of the observer’s movement and therefore compensated for. The compensation phenomenon is known as perceptual constancy (e.g., Gregory, 1963). This mechanism suggests how the distal object is maintained, not how it emerges. Below we consider the relative contribution of this prior compared to that of the signal-receptor interaction to shaping the experience of an external world.

One way to explore this is to look at the modality specific constraints to distal attribution and measure their influence. If modality specific constraints prevent attribution to a distal source then it is reasonable to assume that the process is not amodal in nature but depends on cues that are modality specific. For example, visual-to-tactile SSDs may provide inputs that obey visual laws or principles. Yet, it is also possible that such constraints do not apply to the tactilely sensed information. There is a vibrant discussion addressing this is-

sue of whether the novel emergent perceptual space is defined by either the substituting or substituted modality, or something else altogether (Auvray and Myin, 2009; Deroy and Auvray, 2012; Loomis *et al.*, 2012).

Feedback about the world (the result of an action/perception loop) makes it possible to establish the mapping between self-generated movement and changes in the proximal stimulation. One question is what is altered in our representation of space when information is remapped from one sensory modality to another with SSDs? Using a visual-to-tactile SSD, Epstein *et al.* (1986) delivered patterns of vibrotactile stimulation to the participants' index fingertip. They assessed how the participants attributed the transformations of vibrotactile stimulation as a consequence of self-generated movement. Attribution was measured through asking participants to rate how well a variety of scenarios matched what they experienced. Analysis of the ratings from these scenarios revealed that although the participants became aware of the relationship between self-generated movement and stimulation transformations, they never chose the scenario that corresponded to the real set-up, i.e., the SSD set up. The authors concluded that their participants had not developed the hypothesis of distal sources, that is, the hypothesis that the ultimate cause of their vibrotactile experience was an encounter with an object in an external world. Here we see an example in which no distal cause is assigned probably because participants do not have an assumption that the stimulation is part of an external world.

However, a second study by Auvray *et al.* (2005), using a visual-to-auditory sensory substitution device (the Vibe, see Hanneton *et al.*, 2010) and a similar method to that of Epstein *et al.* (1986) demonstrated that people appeared to be able, under certain conditions, to attribute stimulation to a distal object. In this study, the scenarios were re-written in order to determine which component of distal attribution the participants reached: the existence of a correlation between their movements and the resulting sensory stimulation (coupling), the existence of an object that caused their sensations (object), and the existence of external space produced by this coupling (space). In the conditions in which there were no additional cues to that provided in Epstein *et al.*'s study, participants in Auvray *et al.*'s study similarly failed to reach the hypothesis of a distal cause although participants indicated that they experienced a sensorimotor coupling between their own actions and the sensed information.

In this case, the device rendered the objects in such a way that they could plausibly exist independently of the existence of external space. However, in a second condition Auvray *et al.* gave their participants a cardboard sheet that they could use as an occluder. The cardboard served as an occluder because by moving the sheet vertically participants could interrupt the source of stimulation. As such intermittent calibration of external space occurred because, if the cardboard could interrupt the source of stimulation, it must therefore be

located *behind* the cardboard, i.e., in external space. These two studies demonstrate the importance of (a) the assumption of an external stable world and (b) the ability to interact with that world and account for the perturbations caused by the interactions.

Other studies used both 2D and 3D spatial estimates (White *et al.*, 1970) and egocentric distance estimates (Siegle and Warren, 2010) to specify the external world. As such they invoked the stable world prior, which informed the way users made sense of the SSD input. White *et al.* (1970) observed that SSD users were able to estimate the slant of a tactually sensed visual object. That is, they had access to the third dimension of external space in their novel perceptual space. Siegle and Warren tested whether observers were more veridical in distance estimates when they focused on the proximal patterns of stimulation or on the distal object. Both the study by White *et al.* (1970) and Siegle and Warren (2010) demonstrated improvements in performance when observers assumed an external 3D world. From the work mentioned so far it is possible to see that SSDs offer a unique opportunity to study distal attribution as it emerges. However, as discussed here, the SSD studies also have their limitations.

### 3.3. *The Role of Movement*

What are the cues necessary for the nervous system to assign a distal cause to the proximal stimulation at a receptor surface? Many studies note the importance of movement (e.g., Bach-y-Rita *et al.*, 1969; Haynes *et al.*, 1965; White, 1970): In all SSD studies, active exploration of the environment is necessary to establish the existence of external space. During exploration the organism can determine the mapping between changes in the incoming stimulation and its self-generated movements (Auvray *et al.*, 2005; Bach-y-Rita *et al.*, 1969; Hurley and Noë, 2003; Loomis, 1992; O'Regan and Noë, 2001; von Békésy, 1955; White *et al.*, 1970). Movement also allows us to account for deformations that are derived from self-generated movement or physical displacement. For example, spatial position of physical objects is mapped onto retinal location when objects are sensed visually. If an external object caused the stimulation pattern then eye movements should alter the retinal location of the object; whereas if the input is not linked to a position in external space, then eye-movements should have no impact on the perceived location of the stimulation. For proximal attribution on the other hand, incoming sensory information and your efference information should not be correlated. For example, the locus of pain in your fingertip remains at your fingertip even if you move your finger relative to the rest of your body. Therefore being able to account for changes in feedback about the environment as a consequence of exploring it is essential for externalisation of the proximal stimulation.

### *3.3.1. Space from Eye Movements*

If we take the example of visual processing, the retinotopic coding of the eye might suggest that the location can be read off from the receptor array, yet the retinal image is not the only source of spatial information. Important contributions to the updating of spatial representations come from efferent oculomotor signals (von Helmholtz, 1925), as revealed by experiments in which stimuli are displayed immediately before or during saccades (Hallett and Lighthstone, 1976; Wurtz, 2008). The role of local movement patterns coupled to afferent information could also be derived from the extra-ocular muscle proprioception as was suggested by Sherrington (1918). However, experimental evidence on the function of eye proprioception has remained controversial (Donaldson, 2000), leading to the idea that this signal is primarily used for oculomotor calibration and learning rather than for spatial representation (Lewis and Maler, 2001; Poletti *et al.*, 2013). However, in a first instance the correlation can be used to define stimulation in external co-ordinates. Extra-retinal signals can consequently calibrate this space. Whether the correlation between the extra-ocular muscle movement and the retinal image combined is sufficient to experience 3D space without early calibration from touch is still a contentious issue (e.g., Gori *et al.*, 2010).

When we consider space perception, we usually refer to the process through which humans and other organisms become aware of the relative positions of their own bodies and objects around them. Space perception provides cues, such as depth and distance, which are important for movement and orientation in the environment. At a functional level if we consider that space emerges in the interaction between the organism and its environment (e.g., Terekhov and O'Regan, 2013) then spatial perception is derived from the organism establishing the relationship between action and the resulting stimulation patterns. Furthermore, the ability to perceive different spatial dimensions depends upon the constraints of each sensory modality to move within two or three dimensions. In the case of the latter, what assumptions constrain the space that emerges through SSD-information? How do these assumptions differ from those derived by exploration and perception during unmediated circumstances? To tentatively address these questions let's consider the experience of three-dimensional space during normal development. In a series of intriguing studies into spatial constancy, Bower (1964, 1965, 1966a, b) explored whether infants perceive only the retinal projection or the external object. Using generalization of operant responding to assess discrimination, Bower demonstrated that two-month-old infants see the constant real size and shape of an unfamiliar object despite variation in its projective size and shape produced by altering distance and slant. In addition, consistent with processing of visual objects in adult humans, these infants detect variation in the object's physical size and shape when its angular projection remains constant. This evidence for object

representation in such a young developing nervous system suggests that attribution of the proximal stimulation to a coherent external object is present in pre-crawl infants. Thus, simple eye and head movement appear to be sufficient to establish an external cause of the proximal stimulation patterns.

Similar findings (e.g., Caron *et al.*, 1978; Shuwairi *et al.*, 2007) also highlight some of the behavioural changes (and the corresponding evoked potential changes) that allow somatosensory externalization (Rigato *et al.*, 2014). Rigato *et al.* (2014) provided recent insights into how the developing nervous system establishes the external location of visual and somatosensory objects and re-maps the proximal stimulation (in receptor coordinates) into external coordinates. These authors demonstrated that making the transformation from proximal stimulation to an external source occurs slowly during the first year of life, and involves not only neural response changes but the ability to update for postural changes in the infant's own body representation. Therefore infants require knowledge about effector position changes in order for them to be able to externalise incoming patterns of stimulation.

The knowledge derived from movement, both for distance perception and distal attribution is the corner stone of the sensorimotor tradition (e.g., O'Regan and Noë, 2001). From a more traditional viewpoint, von Helmholtz (1909) also observed its importance: "It is only by voluntarily bringing our organs of sense in various relations to the objects that we learn to be sure as to our judgments of the causes of our sensations. We explain the table as having existence independent of our observation because, at any moment we like, simply by assuming the proper position with respect to it, we can observe it." It should be mentioned that this necessity of a structured correlation between actions and sensations in order to allow distal attribution was nicely anticipated by Condillac (1754). If we were only able to passively receive sensations, we would not understand that these sensations refer to objects that exist in an external world. Indeed, if all our knowledge about the world came from our sensations, and if sensations were just passive modifications of our minds, how could we infer the existence of an external world? For Condillac, our exploratory movements allow us to extract the spatial organisation present in our sensations; it is this spatial organisation of the objects in the world that allows us to consider them as external. In other words, spatialisation and object-hood emerge simultaneously.

It should be noted that the movement necessary to reach distal attribution goes beyond a mere coupling between sensation and actions; it must also reflect a group structure. In other words, the experience of external space cannot be reduced to correctly extracting the existence of a coupling between action and sensation. Indeed, understanding a correlation between self-generated movements and their resulting stimulation involves understanding that *different* actions give rise to *different* sensations. There is a space of displacements

when the *same* actions can give rise to *different* sensations, and when *different* actions can give rise to the *same* sensation. For instance, von Helmholtz (1909) highlighted that actions that can cancel each other from the viewpoint of sensory inputs determine a mathematical group structure, and properties of such a group characterize displacements within a geometrical space. The same sensation obtained by different actions can then be understood as “position” of an object. This was concisely explained by Poincaré (1905) according to who reversibility is the only way to differentiate a change in state from a change in position.

Movement along multiple axes is useful for establishing the existence of external objects. With SSD study results, it is tempting to try to disentangle which movement trajectories are most useful in this process, e.g., translation (changing the position of an object), or rotation (circular movement around a centre point). Comparing the movement type (translation and rotation) across studies demonstrating either distal attribution or a lack thereof, however does not lead to a definitive picture. Both Siegle and Warren (2010) and Epstein *et al.* (1986) allowed observers to translate the sensor through space. In one case the consequent vibrotactile pattern was attributed to an external object (Siegle and Warren, 2010) while in the other, no distal attribution emerged (Epstein *et al.*, 1986). As such we can simply conclude that correlated changes between self-generated movement (even with a camera end-point) and sensory feedback are necessary, without being able to classify the optimal motion type. However, the essential difference between the two studies for this review is that Siegle and Warren’s study involved the prior of an external world whereas Epstein *et al.*’s study did not, suggesting this prior to be an important factor for distal attribution.

### 3.4. *Calibrating and Maintaining External Space*

We have discussed the importance of a prior of a stable world and the role of movement in the emergence of distal attribution. In the following sections we address how our experience of external space is maintained via sensory calibration and spatial maps in non-mediated perception. Although a lot is known about the role of calibration of space for information that is conveyed by functional sensory modalities (e.g., King *et al.*, 1988; Knudsen, 1999; Lee and Groh, 2009), the role that calibration plays in maintaining an experience of external space in SSD environments has been rather neglected. In the following two sections we discuss how calibration operates for perception in non-mediated conditions and then postulate its possible role in SSD-mediated external space.

#### 3.4.1. *The Potential Role of Calibration*

All measurement systems require calibration to become and remain accurate. Our perceptual system is no different. Calibration, in the case of our senses,

usually refers to the process of adaptation as a way of producing environmentally geared behaviour (Mon-Williams and Bingham, 2007). While this is particularly critical in vision given that our visual system has no direct access to attributes such as distance, solidity, and size (Ho *et al.*, 2009), cross-sensory calibration occurs for most senses at one time or another: Visual object attributes must be calibrated, that is, verified through touch; auditory position must be calibrated by vision (e.g., King *et al.*, 2001). In each instance the sense with the least error calibrates its co-occurring counterpart. Indeed, since vision is often distorted (see Gori *et al.*, 2010) haptic feedback may be fundamental in improving visual perception through calibration. Calibration has been recognized as essential for maintaining a coherent representation across the senses (e.g., Wallace *et al.*, 1998). In calibration it is important that the system receives feedback (traditionally an error signal) in order to be able to adapt to the conditions it currently experiences: changes in object information occur as we move through the world, calibration is necessary to link the consequent changes in stimulation patterns to the experienced objects (e.g., Ernst, 2008).

However, calibration may also be fundamental to more gradual processes occurring during development, in which a sensory modality calibrates (or teaches) the others about some properties of the world. For instance, recent studies reinforce the suggestion that the haptic system has a role in calibration of the visual system in judgments of position and size. Both animal and human studies have highlighted the role of cross-sensory calibration in the development of spatial processing (Bergan and Knudsen, 2009; Gori *et al.*, 2011; Zwiers *et al.*, 2001, 2003).

If the nervous system is to make full use of the spatial relations among objects and use these relationships to interact with the environment, then online (re)calibration is essential. Experiments in which animals have been reared with modified auditory or visual inputs have revealed substantial plasticity in the animal's spatial representations (e.g., Knudsen, 1999). If animals are raised with distorted binaural cues (used for sound localisation) locating auditory objects in external space produces compensatory changes that can be induced in the auditory space map (e.g., in ferrets, see King and Parsons, 1999; and barn owls, Gold and Knudsen, 2000). These adjustments in auditory spatial tuning tend to preserve the alignment with the visual representation. In owls, they are brought about by frequency-specific shifts in neuronal tuning to interaural time differences (ITDs) and interaural level differences (ILDs) (Gold and Knudsen, 2000), although, as discussed below, the basis for this adaptive plasticity in humans may be different. It is also important that information coming in from completely different co-ordinate systems can be integrated if co-occurring and segregated if not. Spatial maps are one way to achieve this (see Lee and Groh, 2009).

The human (and other animal) nervous system clearly establishes spatial maps to enable efficient processing of incoming information and allow the stimulation to be mapped from receptor to external space. As such there is a need for calibration across the sensors to verify object positions in external coordinates (for example, see Lee and Groh, 2009; Mancini and Haggard, 2014). From the studies exploring space perception via SSDs, there is almost no consideration given to the role of such calibration in learning to perceive the proximal stimulation as distal and maintain this experience over time (e.g., Auvray *et al.*, 2005; Bach-y-Rita, 1972; Bubic *et al.*, 2010; Epstein *et al.*, 1986; Hanneton *et al.*, 2010; Siegle and Warren, 2010). One could speculate that the SSD experiments require such a long training phase not only due to learning the translation code between substituted and substituting modality but also as they may involve a calibration phase. In such an interpretation the nervous system would calibrate the space by feedback derived from self-generated movement.

When we experience visual or tactile objects, they are generally coded within a specific set of coordinates that allows the nervous system to extract the relative position of the object in space. The nervous system establishes spatial maps at low-level processing sites and the perceptual information is segregated according to these maps throughout the processing hierarchy. Sensorimotor acts involve multiple transformations: When we point to an object, a chain of nested coordinate transformations occurs. These include the retinal map itself (target-to-eye), eye-to-head, head-to-body, and body-to-arm mappings. The complexity of coordinating such simple actions is aided by spatial map alignment. However, in exploring the relationship between distance perception and distal attribution, the role of spatial maps and calibration for updating the perceptual space need to be considered separately.

### *3.4.2. Spatial Maps and the Calibration of External Space*

Before crossmodal spatial calibration can effectively guide interactions in the world the incoming information needs to be organised in a spatial field in which coordinate systems can be compared and integrated. Not only are signals in the environment organised in a spatial pattern but this pattern is mirrored at a neural level for vision, audition, and touch (see King and Moore, 1991). In fact the amount of plasticity evident at the level of the superior colliculus in many animals is a clear indicator that the sensory maps are used to coordinate action and integrate information across the different coordinate systems. The assumed functional use of spatial maps in the superior colliculus is that they are for optimising orientation responses to incoming information. Detecting and orienting to objects in external space is already one step up from the simple process of using online feedback (from movement) to establish the existence of an external space. The question posed here is whether



the documented training required for SSD-users to experience proximal stimulation as distal is required because training and movement allow the user to 're-establish' or recalibrate existing maps. These maps definitely play a role in non-mediated perception with neurophysiological evidence for the existence of spatial maps both in the superior colliculus and ascending to higher cortical regions like A1 (King *et al.*, 1988; Zwiers *et al.*, 2003). Moreover, for estimates of distance there is evidence that calibration is essential (see Smets *et al.*, 1987) for the percept to emerge. We therefore suggest that whether the spatial maps emerge within the interaction or are established at a neural level, they appear to play a role in maintaining the experience of an external 3D world.

### 3.4.3. Calibration and Distance

As was discussed in Section 3.4.1, calibration is important in maintaining the veridical perception of external space and it therefore plays an important role in distal attribution. The process is also important for distance estimates using 2D surfaces. Smets *et al.* (1987) suggested that distance perception (on a 2D surface) relies upon a focus point; i.e., a reference against which to calibrate the estimate. Calibration of space increases the veridicality of the judgment. This is also true for the auditory space where relative distance estimates are more accurate than those derived from absolute cues (Shinn-Cunningham, 2001). In distal attribution, a focus point (also achieved through an interposition of two objects) is important for calibrating external space. Establishing spatial relations among objects and perceivers allows us to extract the different reference frames that can be used to describe external space (see, e.g., Knapen *et al.*, 2010; Poletti *et al.*, 2013).

If we take the example of visual perception, the distance of a visual object cannot be directly sensed (Koenderink *et al.*, 2001). However, perception does not only depend on currently sensed information and sensor position. Experience is also shaped by our previous interactions with the world through which we move, our assumptions about the physical world and underlying principles. For example, the emergence of depth is possible due to the priors we have about the world and the physical laws that constrain the (visually) sensed objects and events therein. We assume the world is stable (Glennerster *et al.*, 2006). One piece of evidence for the strength of our stable world prior is the ability to compensate for changes in an object's size and shape. Perceptual constancies of shape and size are used to compensate for distance in vision are effective because of our prior of a stable world (Glennerster *et al.*, 2006). If we assumed that objects could change their shape or form, then these cues to distance would no longer be valid. There is no origin for depth per se in physical terms although a metric of length is valid. However, a ruler for visual depth, for example, is an internal scale, a compromise between priors about object

constancy and a stable world, likelihood given the retinal projection and the compensation mechanisms provided by our nervous system.

The existence of a distal visual object, as opposed to the proximal pattern of stimulation it affords, can be verified via touch. As such it could be suggested that crossmodal calibration gives the nervous system a direct access to the distal form of objects. To reiterate, distance is an estimate that defines an external world that exists independently of the perceptual processing of the information. Organisms perceive objects in space. Such perception represents the response of the organism to a complex pattern of stimulation, integrated with the remaining effects of past experience with similar objects.

#### **4. Summary**

To highlight what has been learnt about distal attribution through the prism of SSDs, let's explore the factors that appear to be essential to its emergence. First distal attribution depends upon the assumption (or prior) of an external world. Non-separable from this belief of an external world, is the coupling between self-generated movements and the resulting changes in stimulation. This factor is crucial given that if there is no correlation between changes in motor output and changes in proximal patterns of stimulation, then the stimulation will be experienced solely as sensation, not perception (as was made evident in relation to the example of pain). The coupling emerges due to an organism's ability to move and explore the physical world. From the movement patterns the organism can learn about the constraints on external objects such that they can change only in two ways: their position according to the laws of physics or their state according to the laws of nature. In Bayesian terms this second level is the likelihood or currently available stimulation. Using both prior and likelihood the nervous system can then assign a cause, distal or receptor based, given the proximal pattern and its coupling to sensor movement. In this context, distance perception can be seen as a bi-product of the distal attribution process with both phenomena depending on the prior of a stable world (e.g., Glennerster *et al.*, 2006; Knill, 2007). Finally, calibration also plays an important role in maintaining the experience of a stable external world. The ability to assign co-ordinates to the object position in space is possible due to sensory maps which help translate and maintain the input from external world co-ordinates into perceptual co-ordinates.

To summarize, distal attribution of incoming sensory information is influenced by three main parameters: the prior of an external world, the coupling between movements and the resulting changes in stimulation, and calibration. Furthermore, distal attribution has proved crucial for distance perception, in that estimating the co-ordinates of an object in external space requires the object to be perceived as distal, rather than as a proximal pattern of stimulation.

Thus, without the process of distal attribution, perceived distance would not be possible.

### *Acknowledgements*

We would like to thank Irene Fasiello for the invigorating discussions regarding the nature of distal attribution. We would also like to thank Verena Conrad for her input on earlier drafts of this manuscript. Jess Hartcher-O'Brien is supported by a two-year post-doctoral fellowship from the Foundation Fyssen. Malika Auvray was funded by a grant from the Agence Nationale de la Recherche (ANR-11-JSH2-003-1). The DEC supported via ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL.

### **References**

- Auvray, M., Hanne-ton, S., Lenay, C. and O'Regan, J. K. (2005). There is something out there: distal attribution in sensory substitution twenty years later, *J. Integr. Neurosci.* **4**, 505–521.
- Auvray, M., Hanne-ton, S. and O'Regan, J. K. (2007). Learning to perceive with a visuo-auditory substitution system: localisation and object recognition with 'The vOICe', *Perception* **36**, 416–430.
- Auvray, M. and Myin, E. (2009). Perception with compensatory devices: from sensory substitution to sensory motor extension, *Cogn. Sci.* **33**, 1036–1958.
- Bach-y-Rita, P. (1972). *Brain Mechanisms in Sensory Substitution*. Academic Press, London, UK, and New York, NY, USA.
- Bach-y-Rita, P. (2002). Sensory substitution and qualia, in: *Vision and Mind*, A. Noë and E. Thompson (Eds), pp. 497–514. MIT Press, Cambridge, MA, USA.
- Bach-y-Rita, P., Collins, C. C., Saunders, F., White, B. and Scadden, L. (1969). Vision substitution by tactile image projection, *Nature* **221**, 963–964.
- Bergan, J. F. and Knudsen, E. I. (2009). Visual modulation of auditory responses in the owl inferior colliculus, *J. Neurophysiol.* **101**, 2924–2933.
- Bower, T. G. R. (1964). Discrimination of depth in premotor infants, *Psychonom. Sci.* **1**, 368.
- Bower, T. G. R. (1965). Stimulus variables determining space perception in infants, *Science* **149**, 88–89.
- Bower, T. G. R. (1966a). Slant perception and shape constancy in infants, *Science* **151**, 832–834.
- Bower, T. G. R. (1966b). The visual world of infants, *Sci. Am.* **215**, 80–92.
- Bubic, A., Striem-Amit, E. and Amedi, A. (2010). Large-scale brain plasticity following blindness and the use of sensory substitution devices, in: *Multisensory Object Perception in the Primate Brain*, M. J. Naumer and J. Kaiser (Eds), pp. 351–380. Springer, New York, NY, USA.
- Caron, A. J., Caron, R. F. and Carlson, V. R. (1978). Do infants see objects or retinal images? Shape constancy revisited, *Infant Behavior and Development* **1**, 229–243.
- Condillac, E. B. (1754). *Traité des Sensations [Treatise of Sensations]*. Reissued in the series *Corpus des Œuvres de Philosophie en Langue Francaise*. Fayard, Paris, France, 1984.

- Crick, F. and Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature* **375**, 121–123.
- Deroy, O. and Auvray, M. (2012). Reading the world through the skin and ears: a new perspective on sensory substitution, *Front. Theor. Phil. Psychol.* **3**, 457.
- Donaldson, I. (2000). The functions of the proprioceptors of the eye muscles, *Phil. Trans. R. Soc. B Biol. Sci.* **335**, 1685–1754.
- Epstein, W., Hughes, B., Schneider, S. and Bach-y-Rita, P. (1986). Is there anything out there? A study of distal attribution in response to vibrotactile stimulation, *Perception* **15**, 275–284.
- Ernst, M. O. (2008). Multisensory integration: a late bloomer, *Curr. Biol.* **18**, R519–R521.
- Gibson, J. J. (1962). Observations on active touch, *Psychol. Rev.* **69**, 477–490.
- Glennester, A., Tcheang, L., Gilson, S. J., Fitzgibbon, A. W. and Parker, A. J. (2006). Humans ignore motion and stereo cues in favor of a fictional stable world, *Curr. Biol.* **16**, 428–432.
- Gold, J. I. and Knudsen, E. I. (2000). Abnormal auditory experience induces frequency-specific adjustments in unit tuning for binaural localization cues in the optic tectum of juvenile owls, *J. Neurosci.* **20**, 862–877.
- Gori, M., Sandini, G., Martinoli, C. and Burr, D. (2010). Poor haptic orientation discrimination in nonsighted children may reflect disruption of cross-sensory calibration, *Curr. Biol.* **20**, 223–225.
- Gori, M., Sciutti, A., Burr, D. and Sandini, G. (2011). Direct and indirect haptic calibration of visual size judgments, *PLoS ONE* **6**, e25599.
- Graham, N. (1989). *Visual Pattern Analyzers*. Oxford University Press, New York, NY, USA.
- Graziano, M. S., Reiss, L. A. and Gross, C. G. (1999). A neuronal representation of the location of nearby sounds, *Nature* **397**, 428–430.
- Gregory, R. L. (1963). Distortions of visual space as inappropriate constancy scaling, *Nature* **199**, 678–680.
- Guarniero, G. (1974). Experience of tactile vision, *Perception* **3**, 101–104.
- Hallett, P. E. and Lightstone, A. D. (1976). Saccadic eye movements towards stimuli triggered by prior saccades, *Vis. Res.* **16**, 99–106.
- Hanneton, S., Auvray, M. and Durette, B. (2010). The Vibe: a versatile vision-to-audition sensory substitution device, *Appl. Bionics Biomech.* **7**, 269–276.
- Haynes, H., White, B. L. and Held, R. (1965). Visual accommodation in human infants, *Science* **148**, 528–530.
- Hayward, V. (2011). Is there a ‘plenhaptic’ function? *Phil. Trans. R. Soc. B Biol. Sci.* **366**, 3115–3122.
- Ho, Y. X., Serwe, S., Trommershauser, J., Maloney, L. T. and Landy, M. S. (2009). The role of visuohaptic experience in visually perceived depth, *J. Neurophysiol.* **101**, 2789–2801.
- Hurley, S. and Noë, A. (2003). Neural plasticity and consciousness, *Biol. Philos.* **18**, 131–168.
- Katz, D. (1925). *The World of Touch*. Erlbaum, Hillsdale, NJ, USA (translation published 1989, translated by L. E. Kmegeer).
- King, A. J., Hutchings, M. E., Moore, D. R. and Blakemore, C. (1988). Developmental plasticity in the visual and auditory representations in the mammalian superior colliculus, *Nature* **332**, 73–76.
- King, A. J. and Moore, D. R. (1991). Plasticity of auditory maps in the brain, *Trends Neurosci.* **14**, 31–37.
- King, A. J. and Parsons, C. H. (1999). Improved auditory spatial acuity in visually deprived ferrets, *Eur. J. Neurosci.* **11**, 3945–3956.

- King, A. J., Schnupp, J. W. and Doubell, T. P. (2001). The shape of ears to come: dynamic coding of auditory space, *Trends Cogn. Sci.* **5**, 261–270.
- Knapen, T., Rolfs, M., Wexler, M. and Cavanagh, P. (2010). The reference frame of the tilt aftereffect, *J. Vis.* **10**, 1–13.
- Knill, D. C. (2007). Learning Bayesian priors for depth perception, *J. Vis.* **7**, 1–20.
- Knudsen, E. I. (1999). Mechanisms of experience-dependent plasticity in the auditory localization pathway of the barn owl, *J. Comp. Physiol.* **185**, 305–321.
- Koenderink, J. J., van Doorn, A. J., Kappers, A. M. L. and Todd, J. T. (2001). Ambiguity and the “mental eye” in pictorial relief, *Perception* **30**, 431–448.
- Krueger, L. E. (1970). David Katz's *Der Aufbau der Tastwelt* (The World of Touch): a synopsis, *Atten. Percept. Psychophys.* **7**, 337–341.
- Lee, J. and Groh, J. M. (2009). Auditory signals evolve from hybrid- to eye-centered coordinates in the primate superior colliculus, *J. Neurophysiol.* **108**, 227–242.
- Lenay, C., Gapenne, O., Hanneton, S., Marque, C. and Genouëlle, C. (2003). Sensory substitution: limits and perspectives, in: *Touching for Knowing*, Y. Hatwell, A. Streri and E. Gentaz (Eds), pp. 275–292. John Benjamins, Amsterdam, The Netherlands.
- Levy-Tzedek, S., Hanassy, S., Abboud, S., Maidenbaum, S. and Amedi, A. (2012). Fast, accurate reaching movements with a visual-to-auditory sensory substitution device, *Restor. Neurol. Neurosci.* **30**, 313–323.
- Lewis, J. E. and Maler, L. (2001). Neuronal population codes and the perception of object distance in weakly electric fish, *J. Neurosci.* **21**, 2842–2850.
- Loomis, J. M. (1992). Distal attribution and presence, *Presence* **1**, 113–119.
- Loomis, J. M., Klatzky, R. L. and Giudice, N. A. (2012). Sensory substitution of vision: importance of perceptual and cognitive processing, in: *Assistive Technology for Blindness and Low Vision*, R. Manduchi and S. Kurniawan (Eds), pp. 162–191. CRC Press, Boca Raton, FL, USA.
- Mancini, F. and Haggard, P. (2014). Perception: a motion after-effect for voluntary actions, *Curr. Biol.* **24**, R70–R72.
- Merleau-Ponty, M. (1945). *Phénoménologie de la Perception*. Gallimard, Paris, France.
- Mon-Williams, M. and Bingham, G. P. (2007). Calibrating reach distance to visual targets, *J. Exp. Psychol. Hum. Percept. Perform.* **33**, 645–656.
- Moore, D. R. and King, A. J. (1999). Auditory perception: the near and far of sound localization, *Curr. Biol.* **9**, R361–R363.
- O'Regan, J. K. (1992). Solving the “real” mysteries of visual perception: the world as an outside memory, *Can. J. Psychol.* **46**, 461–488.
- O'Regan, J. K. and Noë, A. (2001). A sensorimotor account of vision and visual consciousness, *Behav. Brain Sci.* **24**, 939–973.
- Poggio, G. F. and Fisher, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkeys, *J. Neurophysiol.* **40**, 1392–1405.
- Poincaré, H. (1905). *La Valeur de la Science [The Value of Science]*. Flammarion, Paris, France.
- Poletti, M., Burr, D. C. and Rucci, M. (2013). Optimal multimodal integration in spatial localization, *J. Neurosci.* **33**, 14259–14268.
- Rigato, S., Begum Ali, J., van Velzen, J. and Bremner, A. J. (2014). The neural basis of somatosensory remapping develops in human infancy, *Curr. Biol.* **24**, 1222–1226.
- Segond, H., Weiss, D. and Sampaio, E. (2005). Human spatial navigation via a visuo-tactile sensory substitution system, *Perception* **34**, 1231–1249.

- Sherrington, C. S. (1918). Observation on the sensual role of the proprioceptive nerve supply of the extrinsic ocular muscles, *Brain* **41**, 332–343.
- Shinn-Cunningham, B. G. (2001). Localizing sound in rooms, in: *Proceedings of the ACM/SIGGRAPH and Eurographics Campfire: Acoustic Rendering for Virtual Environments*. Snowbird, UT, USA.
- Shuwairi, S. M., Albert, M. K. and Johnson, S. P. (2007). Discrimination of possible and impossible objects in infancy, *Psychol. Sci.* **18**, 303–307.
- Siegle, J. H. and Warren, W. (2010). Distal attribution and distance perception in sensory substitution, *Perception* **39**, 208–223.
- Smets, G. J. F., Overbeeke, C. J. and Strathman, M. H. (1987). Depth on a flat screen, *Percept. Mot. Skills* **64**, 1023–1034.
- Terekhov, A. and O'Regan, J. K. (2013). Space as an invention of biological organisms. <http://arxiv.org/abs/1308.2124>
- von Békésy, G. (1955). Human skin perception of traveling waves similar to those on the cochlea, *J. Acoust. Soc. Am.* **27**, 830–841. [Reprinted in: G. von Békésy (1960). *Experiments in Hearing*, E. G. Wever (Ed. and Trans.), pp. 547–634. McGraw-Hill, New York, NY, USA.]
- von Helmholtz, H. (1909). *Physiological Optics*. Optical Society of America, Rochester, NY, USA.
- von Helmholtz, H. (1925). *Treatise on Physiological Optics*. Dover, New York, NY, USA.
- Wallace, M. T., Meredith, M. A. and Stein, B. E. (1998). Multisensory integration in the superior colliculus of the alert cat, *J. Neurophysiol.* **80**, 1006–1010.
- White, B. W. (1970). Perceptual findings with the vision-substitution system, *IEEE Trans. Man-Mach. Syst.* **11**, 54–58.
- White, B. W., Saunders, F. A., Scadden, L., Bach-y-Rita, P. and Collins, C. C. (1970). Seeing with the skin, *Percept. Psychophys.* **7**, 23–27.
- Wurtz, R. H. (2008). Neuronal mechanisms of visual stability, *Vis. Res.* **48**, 2070–2089.
- Zwiers, M. P., Van Opstal, A. J. and Cruysberg, J. R. M. (2001). Two-dimensional sound-localization behavior of early-blind humans, *Exp. Brain Res.* **140**, 206–222.
- Zwiers, M. P., Van Opstal, A. J. and Paige, G. D. (2003). Plasticity in human sound localization induced by compressed spatial vision, *Nat. Neurosci.* **6**, 175–181.