



**HAL**  
open science

## Individual differences in sensory substitution

Gabriel Arnold, Jacques Pesnot-Lerousseau, Malika Auvray

► **To cite this version:**

Gabriel Arnold, Jacques Pesnot-Lerousseau, Malika Auvray. Individual differences in sensory substitution. *MULTISENSORY RESEARCH*, 2017, 30 (6), pp.579-600. 10.1163/22134808-00002561 . hal-03172180

**HAL Id: hal-03172180**

**<https://hal.science/hal-03172180>**

Submitted on 4 Dec 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.

# Individual Differences in Sensory Substitution

**Gabriel Arnold, Jacques Pesnot-Lerousseau and Malika Auvray\***

Institut des Systèmes Intelligents et de Robotique, CNRS UMR 7222,  
Université Pierre et Marie Curie, 4 place Jussieu, 75005 Paris, France

Received 20 August 2016; accepted 16 March 2017

---

## **Abstract**

Sensory substitution devices were developed in the context of perceptual rehabilitation and they 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, they can convert visual information into sounds or tactile stimuli. In this article, we review those studies that investigated the individual differences at the behavioural, neural, and phenomenological levels when using a sensory substitution device. We highlight how taking into account individual differences has consequences for the optimization and learning of sensory substitution devices. We also discuss the extent to which these studies allow a better understanding of the experience with sensory substitution devices, and in particular how the resulting experience is not akin to a single sensory modality. Rather, it should be conceived as a multisensory experience, involving both perceptual and cognitive processes, and emerging on each user's pre-existing sensory and cognitive capacities.

## **Keywords**

Sensory substitution, individual differences, rehabilitation technologies, multisensory processes, blindness, sensory modalities

## **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). These 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 (see Auvray and Myin, 2009, for a review). For visual-to-tactile devices, the translation code can be 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 using 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). However, in this article, we will focus on those devices that were designed to compensate for visual impairments. 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 firefighting 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 and Spence, 2008).

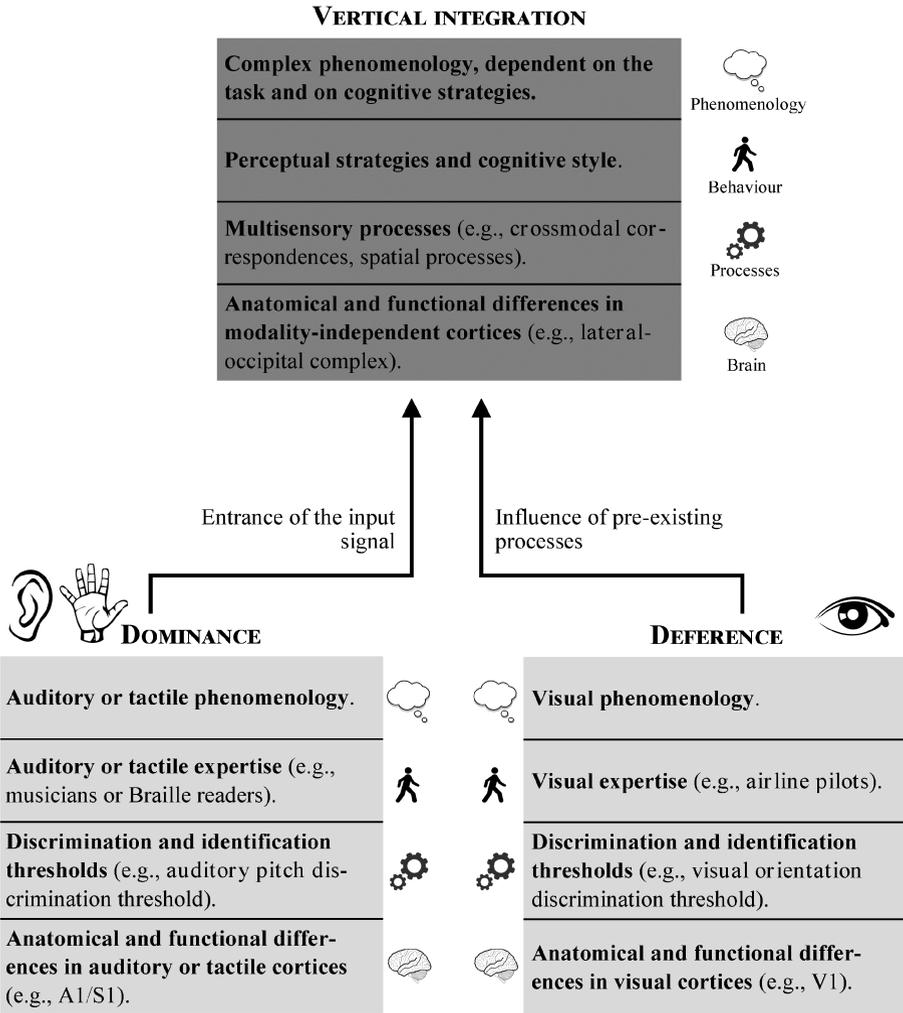
Studies conducted on sensory substitution devices have revealed an important structural and functional plasticity of the central nervous system. As will be seen in detail in this article, the structural and functional plasticity resulting from the use of these devices is highly subject to individual differences. The main findings regarding structural plasticity have shown that the use and practice of visual-to-tactile (Ptito *et al.*, 2005) and visual-to-auditory (e.g., Striem-Amit and Amedi, 2014; for reviews, see Proulx *et al.*, 2014; Stiles and Shimojo, 2015) devices result in increased activation in the blind participants' visual cortex. Regarding functional plasticity, users of such devices can build a perceptual space that possesses visual characteristics although the sensory inputs come from another sensory modality (e.g., Renier *et al.*, 2005). Note, however, that the degree to which the use of a substitution device involves visual processes varies from one user to another (e.g., Auvray *et al.*, 2007). The possibility that the associated subjective experience becomes visual for some users, while remaining tactile or auditory for others, remains insufficiently investigated.

From an applied point of view, the challenge faced by designers consists of finding the best way to provide users with more accurate information usually allowed by vision, such as colour, shape, or distance. This challenge also involves designing devices that are ergonomically adapted to their users (Elli *et al.*, 2014). Improving the design of sensory substitution devices, as will be

further described, will benefit from the identification of individual differences, either between different target populations (e.g., sighted, early-blind, and late-blind people) or within each population, in the different perceptual and spatial processes that are required to learn and acquire a new way of interpreting external information. This will allow the devices to be better adapted to people's specificities in use and learning.

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 and Noë, 2003; Noë, 2004; O'Regan, 2011) according to which perception switches to the substituted modality (vision). The deference thesis 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' (Bachy-Rita *et al.*, 2003). Auvray and Myin (2009), followed by Deroy and Auvray (2012, 2014; see also Farina, 2013) 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, among other processes, both the substituting and the substituted sensory modalities. In addition, sensory substitution should be treated as involving both perceptual and cognitive processes, which take into account integration of the novel information with the existing perceptual-semantic route.

The proposed alternative view has important consequences to understand the use of sensory substitution devices as being dependent on individual capacities (see Fig. 1). First, it 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 strategies, and cognitive styles. Note that these weights might subsequently vary as a function of expertise, appropriation 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



**Figure 1.** Three alternative views (deference, dominance, and vertical integration) on the processes involved in sensory substitution, with their specificities at the behavioural, neural, and phenomenological levels. Illustrative examples are given in parentheses.

that non-musicians could not, consequently giving more weights to auditory processes, and possibly having a predominant auditory experience when using the device.

In this article we will review the individual differences that have been reported regarding the perceptual and spatial processes involved when using sensory substitution devices aiming at compensating visual deficits. Using a sensory substitution device requires learning to extract perceptual information that arrives in the brain *via* an unnatural way and to correctly interpret it. It also

requires correctly organizing this perceptual information spatially, in order to reconstruct the external environment that the device allows one to explore. Regarding perceptual processes, we will first focus on the role of individual abilities in the different sensory modalities, both in the substituted deficient modality (e.g., vision) and in the substituting preserved modality (e.g., audition and touch). Second, we will focus on multisensory processes, particularly on the natural associations that exist between the senses and their implications for optimally converting visual into auditory or tactile stimuli. Regarding spatial processes, we will focus in particular on the reference frames that are involved when integrating multisensory information and when performing navigation tasks with a sensory substitution device. Note that, in this article, we will not discuss more central processes such as learning, attention, or executive function that are also important when using sensory substitution devices. The individual differences reviewed here concern both differences between different populations of users (e.g., sighted and blind people, musicians and non-musicians) and within the same population, taking into account individual perceptual and spatial abilities. We will address individual differences at the behavioural, neural, and phenomenological levels. Finally, we will highlight the implications of individual differences, both for theoretical and applied research.

## **2. Individual Differences in Perceptual Processes**

### *2.1. Modality-Specific Processes*

The theoretical debate opposing the deference and dominance theses has led researchers to investigate whether the use of a sensory substitution device relies on individual abilities both in the substituted and in the substituting sensory modalities. For the substituted modality (i.e., vision), an interesting question is whether perceptual performance with a sensory substitution device is influenced by visual or visuo-spatial abilities. Brown *et al.* (2011) found no influence of individual visual imagery abilities (measured with the Vividness of Visual Imagery Questionnaire, VVIQ; Marks, 1973) on the learning of the vOICE. On the other hand, in a study by Hanneton *et al.* (2015), performance reached by the participants on a task based on the auditory conversion of a visual video game was found to be correlated to the scores the participants obtained in a visual mental rotation test (Vandenberg, 1971). It thus seems that performance with sensory substitution devices is related to only some visuo-spatial abilities. More studies are therefore necessary to investigate the visuo-spatial abilities that are used and those that are not.

Regarding the substituting modality (i.e., audition or touch), differences in the use and appropriation of the devices were found as a function of auditory abilities. Brown *et al.*'s (2011) study revealed that participants with

good musical abilities (in particular discrimination of tonal changes) did not reach better abilities with the device. However, they learned it faster than participants with low musical abilities. This result shows that the use of the device depends on pre-existing auditory abilities. Similar results were found in Pesnot-Lerousseau *et al.*'s (subm.) study. Participants with high auditory abilities (discrimination of intensity, pitch, and duration) were found to have better performance with the vOICE than those with low auditory abilities. Musicians were also found to have a better 'visual' acuity (tested with Snellen E) when using the vOICE than non-musicians (Haigh *et al.*, 2013). Regarding touch, while low-level tactile abilities (discrimination thresholds) have been expected to have an influence on the use of visual-to-tactile substitution devices (e.g., Loomis *et al.*, 2012), no studies have directly investigated this hypothesis. More studies on individual differences in tactile abilities are thus required to estimate the extent to which people differ and the extent to which these differences influence the use of sensory substitution devices.

## 2.2. Multisensory Processing

Individual differences exist at the level of multisensory processes. This has been illustrated, for instance, with cross-modal correspondences, which are associations people make between sensory dimensions across distinct sensory modalities. The existence of such natural associations is particularly relevant for sensory substitution as learning to use sensory substitution devices may be more optimal if the codes that are implemented correspond to these associations (Deroy *et al.*, 2016; Kristjánsson *et al.*, 2016; Stiles and Shimojo, 2015). Cross-modal correspondences have been reported to involve individual differences existing both between blind and sighted people and among sighted people. For instance, Marks (1989) found a cross-modal correspondence between loudness and brightness; however, when looking at individual profiles, this occurred only in around half of the participants. With respect to blind people, Stiles and Shimojo (2015) have reported that, similarly to sighted people but in a weaker way, they intuitively associated (i.e., without training) the vOICE's sounds to tactile textures. On the other hand, Deroy *et al.*'s (2016) study focused on the conversion of the vertical dimension (i.e., visual elevation) into auditory frequency. Their results revealed that sighted participants but not blind participants (neither early-blind, nor late-blind) naturally associate auditory frequency (high-pitched *vs.* low-pitched tone) with the direction of tactile motion (ascending *vs.* descending). One possibility to explain these discrepant results is that blind and sighted people differ in their natural associations only when spatial information is present (e.g., top *vs.* bottom, left *vs.* right). Investigating further the individual differences concerning the sensitivity to cross-modal correspondences between blind and sighted persons and

within each population should therefore allow designing devices more optimally adapted to their target users.

More generally, perception is qualitatively different between blind and sighted people. Due to brain plasticity, the absence of vision allows the recruitment of visual brain areas for tactile tasks (see Sathian and Stilla, 2010, for a review). Such recruitment of supplementary brain areas, combined with the fact that blind people have to rely more on touch, can partly explain why they have a higher tactile acuity than sighted people. Multisensory integration also differs between blind and sighted people. For instance, auditory and tactile stimuli are better integrated in sighted than in blind people (Hötting and Röder, 2004). In addition, blind people inhibit more easily stimuli presented in a task-irrelevant sensory modality (e.g., tactile stimuli for an auditory task) than sighted people (Hötting *et al.*, 2004). The absence of vision, especially in congenitally blind people, has also been reported to have a global impact on spatial cognition, highlighting the important role of vision in the integration of the spatial information that is extracted from different sensory modalities (for a review, see Pasqualotto and Proulx, 2012). The next section will address those individual differences occurring in spatial processes.

### **3. Individual Differences in Spatial Processes**

#### *3.1. Multisensory Reference Frames*

Spatial cognition is subject to inter-individual differences, which can influence the design of both visual-to-tactile sensory substitution devices and navigation aid devices. Individual differences arise both in the spatial reference frames that are adopted and in spatial navigation skills. With respect to the former, it should be noted that individual variability has to be investigated first for each sense, given that spatial reference frames are different from one sensory modality to another (Cohen and Andersen, 2002). Indeed, visual information is initially coded in retinotopic coordinates, 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 a tactile stimulation can be defined according to the stimulated body part, to the entire body, or to the external world (Arnold *et al.*, 2016). These different reference frames can come into conflict, as revealed by the longer time needed to report which hand, the left or the right, has been stimulated first, when the arms are crossed as compared to uncrossed (Shore *et al.*, 2002). The additional time it takes to localize touch in the crossed-hand condition has been interpreted as reflecting spatial transformation from a body-centred to an environment-centred reference frame (Shore *et al.*, 2002).

The diversity of spatial reference frames is even increased when it comes to planning an action, given that spatial information is also coded differently for sensors and action effectors (Cohen and Andersen, 2002). However, spatial transformations allow integrating spatial information coming from the different senses into one common and unified reference frame with the goal of performing actions. This common reference frame is usually an eye-centred reference frame (Cohen and Andersen, 2002). Spatial information coming from vision (Boussaoud and Bremmer, 1999), audition (Zwiers *et al.*, 2003), and touch (Harrar and Harris, 2010) 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 allow the perception of a unified external world rather than multiple worlds (Hartcher-O'Brien and Auvray, 2014).

Thus, when translating visual information into tactile or auditory information, as is the case with sensory substitution devices, the initial spatial information must be coded in a different reference frame. Interestingly, Levy-Tzedek *et al.* (2012) reported that, after training with a visual-to-auditory device, spatial information from one sensory modality can inform movements guided by another sensory modality. However, the spatial transformation of sensory information is not performed in the same way by everyone, as there is an inter-individual variability in the reference frames that are used. For instance, in sighted people, different reference frames are adopted when interpreting letters — stimuli that are usually perceived visually — drawn on the body surface (Arnold *et al.*, 2016). The reference frame that is adopted is egocentric, i.e., centred on the participant's body, for 80% of people, whereas 20% of people adopt a decentred reference frame, i.e., centred on a location external to the body. The individual variability is even stronger as, among egocentric people, some of them adopt a trunk-centred reference frame whereas others adopt a head-centred reference frame. Importantly, the reference frame that is adopted has been reported to be natural rather than a mere cognitive choice. It is worth mentioning here that other studies found differences in the reference frames that are adopted as a function of gender, age, and cultural background (e.g., Goeke *et al.*, 2015). Overall, these results highlight that individual differences in the coding of spatial information have an impact on the way a sensory substitution device is used to explore environment and on the resulting sensory experiences.

### 3.2. Navigation Tasks

An important daily-life spatial task, for which individual differences are particularly salient, is navigation. Indeed, two well-contrasted frames of reference are differently adopted in navigation: an egocentric and an allocentric reference frame (e.g., Tversky, 2003). The egocentric reference frame represents

the location of objects in space relative to the observer's body. On the other hand, the allocentric reference frame represents objects' locations relatively to other objects, and thereby, it involves knowledge of spatial relations that unite the various elements of the environment regardless of the observer's position. The way an itinerary is described orients the reference frame that is adopted (e.g., O'Keefe and Nadel, 1978). In particular, route descriptions are based on an egocentric reference frame. They consist of sequences of instructions specifying the changes in direction necessary to reach the final destination. They are characterized by a serial aspect whose consequence is that if people omit one instruction, they will be unable to reach their destination. By contrast, survey descriptions are based on an allocentric reference frame. They are characterized by a high level of flexibility and have certain superiority over the egocentric representations because they allow reorganization of itineraries such as finding alternative routes or taking shortcuts.

The construction of an allocentric representation requires full integration of the various paths in a geographic perspective using the cardinal directions (Klatzky, 1998). This construction is, in sighted people, greatly facilitated by the use of the visual modality, as it allows a multiple and simultaneous perception of spatial elements, even if they are distant. As a consequence, differences in the reference frame adopted between blind and sighted persons have been reported in many studies. For instance, Noordzij *et al.* (2006) used an environment that was verbally described either from a route description (egocentric frame of reference) or from a survey description (allocentric frame of reference). Their results revealed that blind people's mental representations share the same metric properties as those constructed by sighted people. However, blind people constructed a more robust spatial representation from a route description, whereas sighted people performed better with a survey description. Similarly, Rieser *et al.* (1986) found that when a navigation task requires a change of perspective or spatial inferences (that rely on allocentric reference frames) congenitally blind people do not perform as well as sighted and late-blind individuals. Visual experience also plays a role in the reference frame that is preferentially adopted when both egocentric and allocentric reference frames are possible to solve a spatial memory task (Pasqualotto *et al.*, 2013). Sighted and late-blind participants prefer to adopt an allocentric reference frame, whereas congenitally blind participants, who never experienced vision, prefer to adopt an egocentric reference frame (see Gallay *et al.*, 2013, for a review).

It should be noted that the use of an egocentric perspective is sufficient to succeed in some spatial tasks, for which sighted and blind people can thus obtain comparable performance. To give an example, Loomis *et al.* (1993) compared spatial memory performance in blind and sighted participants. The participants in this study were accompanied along a straight or curved path,

and they were subsequently asked to reproduce it. No significant difference in performance was found between blind and sighted people. This result suggests that visual impairment does not entail the estimation of directions and distances between elements of an environment that has been physically explored. This can be due to the fact that the reproduction of a physically experienced itinerary is based on kinaesthetic and vestibular memory, which involves an egocentric frame of reference. However, when the tasks involve the encoding of spatial information that has not been physically explored, or when they involve spatial inferences or changes in perspective, differences appear between blind and sighted people (e.g., see Rieser *et al.*, 1986).

As we have seen, there is an important inter-individual variability in the reference frames that are adopted when interpreting spatial information, and blind people, compared to sighted people, have a spontaneous tendency to adopt an egocentric perspective as opposed to an allocentric one. These individual differences are important to take into account when designing personalized navigation aid devices. In particular, two options are possible; the first one consists in implementing different reference frames so that people that have difficulties with the allocentric reference frame can use the egocentric one. The second one consists in favouring the development of allocentric reference frames in those persons that have difficulties with it. Regarding this option, Millar (1994) hypothesised that it is the perception of distal elements that allows encoding spatial relations between elements of the environment, and therefore allows the construction of an allocentric reference frame. Thus, such access to distal cues, which can be converted into auditory or tactile information, could be implemented as an option in navigation aid devices. Relevant to this hypothesis, Pasqualotto and Esenkaya (2016) have shown that blind-folded sighted participants are able to elaborate an allocentric representation of a spatial scene when exploring the scene with the vOICE, in the same way as when the scene is explored visually. If similar results were to be found in blind participants, this would provide arguments for the hypothesis that sensory substitution devices can allow blind people to form allocentric spatial representations of their environment.

Finally, the effect of aging — one important source of individual differences in addition to being a cause of visual deficits — has scarcely been investigated. The only study conducted thus far investigating age-related differences when using a sensory-substitution device, found that older adults were less successful in completing a navigation task than younger adults (Levy-Tzedek *et al.*, 2016). This result may reflect a combination of perceptual and cognitive decline with age, and it suggests that older adults could use sensory substitution to compensate visual deficiency. However, significant efforts are necessary for older people to learn how to use the device.

#### 4. Individual Differences at the Phenomenological Level

A few studies have investigated the phenomenology (or subjective experience) associated with the use of a sensory substitution device. Some of these studies highlighted the visual-like character of the experience with the device. For example, Guarniero, a blind philosophy student, who extensively used the Tactile Visual Substitution System, described: “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). In a similar vein, a more recent study found a correlation between the recruitment of the occipital cortex and reports of new experiences of flashes in trained late-blind users of visual-to-tactile sensory substitution devices (Ortiz *et al.*, 2011). However, such changes are not documented in congenitally blind individuals, and thus the visual character of these induced flashes or lights still remains to be investigated to better understand whether the experience is perceptually visual, due to imagination, or even due to residual vision.

Subjective changes have also been documented with visual-to-auditory devices. For instance, in a series of interviews, two blind users of the vOICe reported visual experiences, involving colour, but this phenomenology was limited to certain kinds of objects (Ward and Meijer, 2010). According to these authors, training with the vOICe could lead to the occurrence of visual images, comparable to the synesthetic experiences enjoyed by coloured-hearing synaesthetes. However, these two users had an acquired blindness, and it cannot be ruled out that they associated the sounds to visual images from their memory. Thus, we remain unsure of whether the visual phenomenology is truly perceptual, or if it is just given by memory, or even by imagination, and then merely added to a set of non-visual qualities obtained when using the device (see Auvray and Farina, 2017, for a discussion).

Subjective reports were also gathered in blindfolded sighted participants. In their study, Auvray *et al.* (2007) trained six persons with the vOICe for fifteen hours. The participants were given questionnaires about their qualitative experience while using the device. They were asked which sensory modality their experience most resembled for localization and recognition tasks. The replies varied inter-individually and as a function of the performed task. Localization tasks were more likely to be apprehended either as giving rise to visual experiences or as belonging to a new sense, whereas recognition tasks were judged as more closely resembling audition. Two participants mentioned a resemblance with the tactile modality. In addition, the participants

often emphasised that they simply had the feeling of mastering a new tool. Thus, in this study, while the participants were able to perceive visual properties when they are converted into sound, they did not strictly speaking have a visual phenomenal experience or an auditory one. Rather, the participants' phenomenal experience was task-dependent and its tool-like nature was emphasized.

With the same device, the vOICE, in Pesnot-Lerousseau *et al.*'s (subm.) study, blindfolded sighted participants were trained at localization and recognition tasks. Prior to training, and after training in each of the tasks, they were questioned about their subjective experience by means of Likert scales. As in Auvray *et al.*'s (2007) study, participants had a different phenomenology as a function of the task. The overall phenomenology remained auditory, which is likely due to the fact that the stimuli consisted of sounds. However, in the object recognition task, participants additionally had 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 differences in the phenomenologies as a function of the task — which probably result from the fact that several different processes are involved — are 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 that does not differ with the task, which suggests a lower plasticity in the involved processes. As their auditory inputs are poorer, it could be that they are less able to disentangle the different auditory dimensions and consequently to analyse them independently.

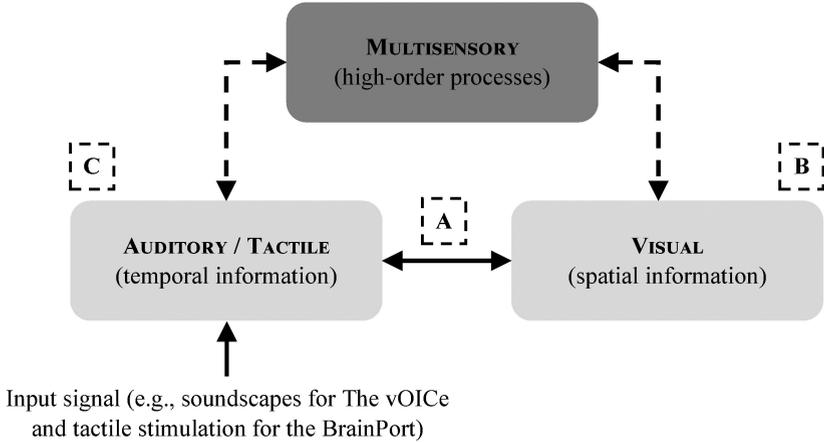
It is not straightforward to know what can be inferred exactly from these reports, as the evidence remains fragile. In blind persons, the reports are compatible with the resurgence of colour or texture memories for familiar objects. Similarly, it should be further investigated what type of visual imagery exactly is at stake, as it might correspond to a mere reminiscence of cross-modal imagery (e.g., Spence and Deroy, 2013), especially in blind people. This survey might also question participants' abilities to establish classes of similarities between their experiences. Nonetheless, what can be inferred from the results is that subjective experience with the device depends both on the task and on individual perceptual abilities.

## 5. Individual Differences at the Neural Level and Proposed Hypothesis of Two Different Processing Pathways

Individual differences at the neural level have been investigated both with visual-to-tactile and visual-to-auditory sensory substitution devices. Regarding visual-to-tactile devices, many studies revealed differences in the activation of the visual cortex in blind and sighted people after training (for reviews, see Poirier *et al.*, 2007; Proulx *et al.*, 2016). For instance, Ptito *et al.* (2005) investigated changes in regional cerebral blood flow (rCBF) using positron-emission tomography (PET) in congenitally blind people and sighted controls. The participants were required to learn to detect the orientation of visual stimuli presented through the device during one hour a day over one week. The blind and sighted participants reached an equal level of performance after training. No task-evoked rCBF increase was detected in the visual cortex of either group before training. After practice, whereas tongue stimulation significantly led to an increase in rCBF in the visual cortex of the blind participants, such activation was not found in the sighted controls. According to the authors, in the case of congenital blindness, training with a visual-to-tactile substitution device induces a rerouting of tactile information to the visual cortex. It should, however, be mentioned that increased activation in the visual cortex of blind participants does not necessarily mean that the corresponding perception is visual.

To test this hypothesis, Kupers *et al.* (2006) used transcranial magnetic stimulation (TMS) of blind and blindfolded sighted participants' visual cortex before and after training with a visual-to-tactile substitution device. Before training, when TMS was applied over their visual cortex, the participants did not report any subjective tactile sensation (only phosphenes were reported by sighted participants). After training, some of the blind participants (three out of eight early-blind participants and one out of five late-blind participants) reported tactile sensations when TMS was applied over the occipital cortex, whereas no such sensations were reported by sighted participants. The authors concluded that the subjective character of the percept would depend on the stimulated sensory channel and not on the activated cortex. In other words, the subjective experience associated with an activity in the visual cortex after training was tactile and not only visual.

Regarding visual-to-auditory devices, in a vast majority of cases, training participants results in an increased activation of their visual areas (for reviews, see Poirier *et al.*, 2007; Proulx *et al.*, 2014, 2016). However, large inter-individual differences have been reported, in particular between early-blind, late-blind, and sighted participants. Compared to late-blind and sighted participants, early-blind participants have stronger activations of their visual cortices (Amedi *et al.*, 2007). More precisely, their visual cortices are more



**Figure 2.** Illustration of the two neural streams described by Poirier *et al.* (2007). Early-blind participants use the cross-modal pathway (solid line). Sighted participants use the mental-imagery pathway (dashed line). We propose three hypotheses to explain this difference. (A) Blind participants already have enhanced cross-modal connections before training with the device. (B) Sensory deprivation allows the recruitment of visual areas. (C) Auditory or tactile abilities determine the recruitment of visual processes. As blind participants have better low-level auditory and tactile capacities, the recruitment of visual areas is facilitated.

modulated by training, and they show a stronger functional connectivity between visual and frontal areas (Murphy *et al.*, 2016). In addition, early-blind people are better at using the device in orientation identification tasks (Stiles *et al.*, 2015) and in ‘visual’ acuity tasks (Striem-Amit *et al.*, 2012). To account for these differences, Poirier *et al.* (2007) proposed a two-stream model. According to this model, early-blind participants would recruit a cross-modal pathway, consisting of a direct stream between auditory cortices and visual ones. On the other hand, late-blind and blindfolded sighted participants, due to their past or current possession of the visual sense, would activate a mental-imagery pathway. Sighted participants, even when blindfolded, would process the information in higher order multimodal cortices and then propagate the signal into visual cortices. This multimodal pathway corresponds to the transduction of the auditory signal not only into visual images, but also potentially into spatial images, that is, a multimodal signal. In this section, three hypotheses are proposed to explain how and why these two pathways can be observed (see Fig. 2).

First, before training, cross-modal connections between visual and auditory cortices may be stronger in early-blind than in late-blind and sighted people. This hypothesis makes sense in the field of the metamodal brain theory (Heimler *et al.*, 2015; Pascual-Leone and Hamilton, 2001; Proulx *et al.*, 2014). If the brain is organised as a task-dependent rather than as a sensory-dependent

system, the ‘visual’ areas are not literally visual. They are involved in spatial processes but their inputs are not restricted to the retina. In early-blind participants, auditory and tactile perceptions are the main source of spatial information. Therefore, their visual cortices may already be strongly linked to auditory and somatosensory cortices, as they are the main provider of useful information. When learning to use the device, blind participants, accustomed to process auditory input with their visual (spatial) cortex, would recruit the cross-modal pathway. However, contrary to this hypothesis, a functional connectivity analysis made by Murphy *et al.* (2016) showed no difference in the auditory–visual cortex connectivity between blind and sighted people.

A second hypothesis is that sensory deprivation and the absence of visual input make visual areas more available to process new signals. As the retina does not send inputs to the cortex, blind participants may recruit those areas more easily. A privation study by Proulx *et al.* (2008) tested this hypothesis. Three sighted participants were blindfolded during three weeks. However, no major differences were found between deprived and control participants. Furthermore, a similar privation study with magnetoencephalography revealed that privation has an effect on the auditory cortex activity only when combined with learning (Pollok *et al.*, 2005). Privation alone did not produce changes. Thus, even if privation induces profound changes in the organisation of the brain (Bavelier and Neville, 2002), it appears that privation alone cannot explain the difference between blind and sighted people.

Finally, auditory capacities may have a role in the observed differences between sighted, early-blind, and late-blind people. Access to the cross-modal pathway may be conditioned by the quality of auditory inputs. As early-blind people have an enhanced auditory sensitivity (Wan *et al.*, 2010), the signal may be sufficiently accurate to be directly processed by visual cortices. Two clues in the literature favour this hypothesis. First, the capacity to use a visual-to-auditory device is theoretically linked to auditory capacities. As shown in Section 2.1, musicians have a better “‘visual’ acuity when using the vOICE (Haigh *et al.*, 2013). Secondly, early-blind people have better low-level auditory capacities than the rest of the population (Wan *et al.*, 2010) and they proved to be better at learning how to use the device (Stiles *et al.*, 2015). This hypothesis makes sense with the two-stream theory (Poirier *et al.*, 2007). It is because blind people’s auditory processes are accurate and because the signal is well discriminated that they have access to another type of treatment, a cross-modal one.

These three hypotheses are attempts to account for the observed differences between early-blind, late-blind, and sighted participants, as well as for the existence of two streams of processes. If prior cross-modal pathways and sensory deprivation have already been studied, the hypothesis of different auditory capacities has been evoked but not directly investigated yet. We believe that

further research is needed to better understand the implication of audition in the use of visual-to-auditory devices, and more generally to understand the role of the input modality in sensory substitution.

## 6. Conclusion and Implications

The experiments conducted with sensory substitution devices have highlighted numerous individual differences in perceptual and spatial processes, at the behavioural, neural, and phenomenological levels. These individual differences have consequences both for the design of sensory substitution devices and for the understanding of the processes involved in a sensory substitution device use.

### 6.1. Implications for the Design of Sensory Substitution Devices

Although sensory substitution appears a promising technology to compensate visual impairments, they remain barely used by visually impaired people, however, either because their design is not optimal, or because their use involves too many constraints and cognitive effort (Elli *et al.*, 2014). Taking into account individual differences will allow an optimal appropriation of these devices. First, the choice of the conversion code will benefit from the identification of the cross-modal correspondences that may exist in the target population. For instance, the existence of an audio-tactile correspondence between pitch and tactile elevation in sighted but not in blind people (Deroy *et al.*, 2016) questions the coding of spatial elevation by auditory frequency that is used in several visual-to-auditory devices. Second, 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). Note that the way each person uses the device (e.g., attributing tactile stimuli to either external objects as in vision or to proximal stimulations as in touch) may be characterized by the adoption of different reference frames, and the preference for one or another may in turn be an indicator of the perceptual experience that is associated with the use of the device.

In addition, identifying different perceptual strategies associated with 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 for 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 (Elli *et al.*, 2014). For instance, for people that have a visual use only, learning could be oriented toward integrating a real 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 that have 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, when designing learning protocols for using sensory substitution, more central sources of individual differences should be taken into account. For instance, aging has been shown to have an impact both on peripheral factors such as perceptual acuity and on central factors such as learning, attention, spatio-temporal integration, and memory (West, 1996). As a consequence, the use of visual-to-tactile sensory substitution in older adults is challenging (Levy-Tzedek *et al.*, 2016) and requires adapting learning protocols to this specific population.

## *6.2. Implications for the Conceptual Understanding of Sensory Substitution*

The data collected on individual differences helps making advances in the understanding of the processes involved when using a sensory substitution device. At the sensory level, it was found that people with good musical abilities (Brown *et al.*, 2011) or good auditory abilities (Pesnot-Lerousseau *et al.*, submitted) have better performance when using visual-to-auditory substitution devices. People with high visuo-spatial abilities have also better performance with visual-to-auditory devices (Hanneton *et al.*, 2015; but see Brown *et al.*, 2011). These results suggest that performance with the device relies on users' individual sensory abilities and that experience takes root both in the substituting and the substituted sensory modalities. The data collected on the subjective experiences, even if they remain too sparse to make strong claims, reveal that they are subject to inter- and intra-individual variability. The phenomenology associated with the use of the device appears to depend both on the task and on auditory abilities. 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. Finally, at the neural level, differences were observed between early-blind, late-blind, and sighted participants, which led to the hypothesis of two streams of processes, one linked to the substituting and substituted modalities and a cross-modal one. Overall, these results are in line with the view that sensory substitution devices are vertically integrated on pre-existing capacities (Deroy and Auvray, 2012, 2014) that involve cognitive and perceptual processes belonging both to the substituting and to the substituted sensory modalities.

## Acknowledgements

This work was supported by the Mission pour l'Interdisciplinarité (CNRS, Auton, Sublima grant) and the labex SMART (ANR-11-LABX-65). GA is funded by a grant from the European Research Council (Wearhap, FP7, Nr. 601165). We thank Mirko Farina for his useful comments on a previous version of the manuscript.

## References

- Amedi, A., Stern, W. M., Camprodon, J. A., Bermpohl, F., Merabet, L., Rotman, S., Hemond, C., Meijer, P. and Pascual-Leone, A. (2007). Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex, *Nat. Neurosci.* **10**, 687–689.
- Arnold, G., Spence, C. and Auvray, M. (2016). Taking someone else's spatial perspective: natural stance or effortful decentring? *Cognition* **148**, 27–33.
- Auvray, M. and Farina, M. (2017). Patrolling the boundaries of synaesthesia: a critical appraisal of transient and artificially-acquired forms of synaesthetic experiences, in: *Sensory Blendings: New Essays on Synaesthesia*, O. Deroy (Ed.). Oxford University Press, Oxford, UK. (Publication scheduled for April 2017).
- Auvray, M. and Myin, E. (2009). Perception with compensatory devices. From sensory substitution to sensorimotor extension, *Cogn. Sci.* **33**, 1036–1058.
- Auvray, M., Hanneton, 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.
- Bach-y-Rita, P., Collins, C. C., Saunders, F. A., White, B. and Scadden, L. (1969). Vision substitution by tactile image projection, *Nature* **221**(5184), 963–964.
- Bach-y-Rita, P., Tyler, M. E. and Kaczmarek, K. A. (2003). Seeing with the brain, *Int. J. Hum-Comput. Interact.* **2**, 285–295.
- Bavelier, D. and Neville, H. J. (2002). Cross-modal plasticity: where and how? *Nat. Rev. Neurosci.* **3**, 443–452.
- Block, N. (2003). Tactile sensation via spatial perception, *Trends Cogn. Sci.* **7**, 285–286.
- Bologna, G., Deville, B. and Pun, T. (2009). On the use of the auditory pathway to represent image scenes in real-time, *Neurocomputing* **72**, 839–849.
- Boussaoud, D. and Bremner, F. (1999). Gaze effects in the cerebral cortex: reference frames for space coding and action, *Exp. Brain Res.* **128**, 170–180.
- Brown, D. J., Macpherson, T. and Ward, J. (2011). Seeing with sound? Exploring different characteristics of a visual-to-auditory sensory substitution device, *Perception* **40**, 1120–1135.
- Cohen, Y. E. and Andersen, R. A. (2002). A common reference frame for movement plans in the posterior parietal cortex, *Nat. Rev. Neurosci.* **3**, 553–562.
- Danilov, Y., Tyler, M., Skinner, K., Hogle, R. and Bach-y-Rita, P. (2007). Efficacy of electro-tactile vestibular substitution in patients with peripheral and central vestibular loss, *J. Vestib. Res.* **17**, 119–130.
- Deroy, O. and Auvray, M. (2012). Reading the world through the skin and ears: a new perspective on sensory substitution, *Fron. Psychol.* **3**, 457. DOI:10.3389/fpsyg.2012.00457.

- Deroy, O. and Auvray, M. (2014). A crossmodal perspective on sensory substitution, in: *Perception and Its Modalities*, S. Biggs, M. Matthen and D. Stokes (Eds), pp. 327–349. Oxford University Press, Oxford, UK.
- Deroy, O., Fasiello, I., Hayward, V. and Auvray, M. (2016). Differentiated audio-tactile correspondences in sighted and blind individuals, *J. Exp. Psychol. Hum. Percept. Perform.* **42**, 1204–1214.
- Diot, B., Halavackova, P., Demongeot, J. and Vuillerme, N. (2014). Sensory substitution for balance control using a vestibular-to-tactile device, *Multisens. Res.* **27**, 313–336.
- Elli, G. V., Benetti, S. and Collignon, O. (2014). Is there a future for sensory substitution outside academic laboratories? *Multisens. Res.* **27**, 271–291.
- Farcy, R., Leroux, R., Jucha, A., Damaschini, R., Grégoire, C. and Zogaghi, A. (2006). Electronic travel aids and electronic orientation aids for blind people: technical, rehabilitation and everyday life points of view, in: *Proceedings of the Conference and Workshop on Assistive Technology for Vision and Hearing Impairment*, M. A. Hersch (Ed.).
- Farina, M. (2013). Neither touch nor vision: sensory substitution as artificial synaesthesia? *Biol. Philos.* **28**, 639–655.
- Gallay, M., Denis, M. and Auvray, M. (2013). Navigation assistance for blind pedestrians: guidelines for the design of devices and implications for spatial cognition, in: *Representing Space in Cognition: Interrelations of Behaviour, Language, and Formal Models*, T. Tenbrink, J. Wiener and C. Claramunt (Eds). Oxford University Press, Oxford, UK.
- Goeke, C., Kornpetpanee, S., Köster, M., Fernández-Revelles, A., Gramann, K. and König, P. (2015). Cultural background shapes spatial reference frame proclivity, *Sci. Rep.* **5**, 11426. DOI:10.1038/srep11426.
- Guarniero, G. (1974). Experience of tactile vision, *Perception* **3**, 101–104.
- 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.
- Hanneton, S., Herquel, P. and Auvray, M. (2015). Intermodal recoding of a video game: learning to process signals for motion perception in a pure auditory environment, *Int. J. Adapt. Contr. Signal Proc.* **29**, 1475–1483.
- Harrar, V. and Harris, L. R. (2010). Touch used to guide action is partially coded in a visual reference frame, *Exp. Brain Res.* **203**, 615–620.
- Hartcher-O'Brien, J. and Auvray, M. (2014). The process of distal attribution illuminated through studies of sensory substitution, *Multisens. Res.* **27**, 421–441.
- Heimler, B., Striem-Amit, E. and Amedi, A. (2015). Origins of task-specific sensory-independent organization in the visual and auditory brain: neuroscience evidence, open questions and clinical implications, *Curr. Opin. Neurobiol.* **35**, 169–177.
- Ho, C. and Spence, C. (2008). *The Multisensory Driver: Implications for Ergonomic Car Interface Design*. Ashgate Publishing, Aldershot, UK.
- Hötting, K. and Röder, B. (2004). Hearing cheats touch, but less in congenitally blind than in sighted individuals, *Psychol. Sci.* **15**, 60–64.
- Hötting, K., Rösler, F. and Röder, B. (2004). Altered auditory-tactile interactions in congenitally blind humans: an event-related potential study, *Exp. Brain Res.* **159**, 370–381.
- Humphrey, N. (1992). *A History of the Mind*. Simon and Schuster, New York, NY, USA.

- Hurley, S. and Noë, A. (2003). Neural plasticity and consciousness, *Biol. Philos.* **18**, 131–168.
- Jones, L. A., Kunkel, J. and Piatetski, E. (2009). Vibrotactile pattern recognition on the arm and back, *Perception* **38**, 52–68.
- Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: definitions, distinctions, and interconnections, in: *Spatial Cognition I*, C. Freksa, C. Habel and K. F. Wender (Eds), pp. 1–18. Springer, New York, NY, USA.
- Kristjánsson, Á., Moldoveanu, A., Jóhannesson, Ó., Balan, O., Spagnol, S., Valgeirsdóttir, V. V. and Unnpórrsson, R. (2016). Designing sensory-substitution devices: principles, pitfalls and potential, *Restor. Neurol. Neurosci.* **34**, 769–787.
- Kupers, R., Fumal, A., Maertens de Noordhout, A., Gjedde, A., Schoenen, J. and Ptito, M. (2006). Transcranial magnetic stimulation of the visual cortex induces somatotopically organized qualia in blind subjects, *Proc. Natl Acad. Sci.* **35**, 13256–13260.
- Levy-Tzedek, S., Novick, I., Arbel, R., Abboud, S., Maidenbaum, S., Vaadia, E. and Amedi, A. (2012). Cross-sensory transfer of sensory-motor information: visuomotor learning affects performance on an audiomotor task, using sensory-substitution, *Sci. Rep.* **2**, 949. DOI:10.1038/srep00949.
- Levy-Tzedek, S., Riemer, D. and Amedi, A. (2014). Color improves “visual” acuity *via* sound, *Front. Neurosci.* **8**, 358. DOI:10.3389/fnins.2014.00358.
- Levy-Tzedek, S., Maidenbaum, S., Amedi, A. and Lackner, J. (2016). Aging and sensory substitution in a virtual navigation task, *PLoS One* **11**, e0151593. DOI:10.1371/journal.pone.0151593.
- Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W. and Fry, P. A. (1993). Nonvisual navigation by blind and sighted: assessment of path integration ability, *J. Exp. Psychol. Gen.* **122**, 73–91.
- 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.
- Maidenbaum, S., Hanassy, S., Abboud, S., Buchs, G., Chebat, D. R., Levy-Tzedek, S. and Amedi, A. (2014). The “EyeCane”, a new electronic travel aid for the blind: technology, behavior and Swift learning, *Restor. Neurol. Neurosci.* **32**, 813–824.
- Marks, D. F. (1973). Visual imagery in the recall of pictures, *Br. J. Psychol.* **64**, 17–24.
- Marks, L. E. (1989). On cross-modal similarity: the perceptual structure of pitch, loudness, and brightness, *J. Exp. Psychol. Hum. Percept. Perform.* **15**, 586–602.
- Meijer, P. B. L. (1992). An experimental system for auditory image representations, *IEEE Trans. Biomed. Eng.* **39**, 112–121.
- Millar, S. (1994). *Understanding and Representing Space: Theory and Evidence From Studies With Blind and Sighted Children*. Clarendon Press, Oxford, UK.
- Murphy, M. C., Nau, A. C., Fisher, C., Kim, S. G., Schuman, J. S. and Chan, K. C. (2016). Top-down influence on the visual cortex of the blind during sensory substitution, *NeuroImage* **125**, 932–940.
- Noë, A. (2004). *Action in Perception*. MIT Press, Cambridge, MA, USA.
- Noordzij, M. L., Zuidhoek, S. and Postma, A. (2006). The influence of visual experience on the ability to form spatial mental models based on route and survey descriptions, *Cognition* **100**, 321–342.

- O'Keefe, J. and Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford University Press, London, UK.
- O'Regan, J. K. (2011). *Why Red Doesn't Sound Like a Bell. Understanding the Feel of Consciousness*. Oxford University Press, Oxford, UK.
- Ortiz, T., Poch, J., Santos, J. M., Requena, C., Martínez, A. M., Ortiz-Terán, L., Turrero, A., Barcia, J., Nogales, R., Calvo, A., Martínez, J. M., Córdoba, J. L. and Pascual-Leone, A. (2011). Recruitment of occipital cortex during sensory substitution training linked to subjective experience of seeing in people with blindness, *PLoS One* **6**, e23264. DOI:10.1371/journal.pone.0023264.
- Pascual-Leone, A. and Hamilton, R. (2001). The metamodal organization of the brain, *Prog. Brain Res.* **134**, 427–445.
- Pasqualotto, A. and Esenkaya, T. (2016). Sensory substitution: the spatial updating of auditory scenes “mimics” the spatial updating of visual scenes, *Front. Behav. Neurosci.* **10**, 79. DOI:10.3389/fnbeh.2016.00079.
- Pasqualotto, A. and Proulx, M. J. (2012). The role of visual experience for the neural basis of spatial cognition, *Neurosci. Biobehav. Rev.* **36**, 1179–1187.
- Pasqualotto, A., Spiller, M. J., Jansari, A. S. and Proulx, M. J. (2013). Visual experience facilitates allocentric spatial representation, *Behav. Brain Res.* **236**, 175–179.
- Pesnot-Lerousseau, J., Arnold, G. and Auvray, M. (subm.). Training induced plasticity with visual-to-auditory sensory substitution devices.
- Poirier, C., De Volder, A. G. and Scheiber, C. (2007). What neuroimaging tells us about sensory substitution, *Neurosci. Biobehav. Rev.* **31**, 1064–1070.
- Pollok, B., Schnitzler, I., Stoerig, P., Mierdorf, T. and Schnitzler, A. (2005). Image-to-sound conversion: experience-induced plasticity in auditory cortex of blindfolded adults, *Exp. Brain Res.* **167**, 287–291.
- Prinz, J. (2006). Putting the brakes on enactive perception, *Psyche* **12**, 1–19.
- Proulx, M. J., Stoerig, P., Ludowig, E. and 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. DOI:10.1371/journal.pone.0001840.
- Proulx, M. J., Brown, D. J., Pasqualotto, A. and Meijer, P. (2014). Multisensory perceptual learning and sensory substitution, *Neurosci. Biobehav. Rev.* **41**, 16–25.
- Proulx, M. J., Gwinnutt, J., Dell'Erba, S., Levy-Tzedek, S., de Sousa, A. A. and Brown, D. J. (2016). Other ways of seeing: from behavior to neural mechanisms in the online “visual” control of action with sensory substitution, *Restor. Neurol. Neurosci.* **34**, 29–44.
- Ptito, M., Moesgaard, S. M., Gjedde, A. and Kupers, R. (2005). Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind, *Brain* **128**, 606–614.
- Renier, L., Laloyaux, C., Collignon, O., Tranduy, D., Vanlierde, A., Bruyer, R. and De Volder, A. G. (2005). The Ponzo illusion using auditory substitution of vision in sighted and early blind subjects, *Perception* **34**, 857–867.
- Rieser, J. J., Gum, D. A. and Hill, E. W. (1986). Sensitivity to perspective structure while walking without vision, *Perception* **15**, 173–188.
- Sathian, K. and Stilla, R. (2010). Cross-modal plasticity of tactile perception in blindness, *Restor. Neurol. Neurosci.* **28**, 271–281.
- Shore, D. I., Spry, E. and Spence, C. (2002). Confusing the mind by crossing the hands, *Cogn. Brain Res.* **14**, 153–163.

- Spence, C. and Deroy, O. (2013). Crossmodal mental imagery, in: *Multisensory Imagery*, S. Lacey and R. Lawson (Eds), pp. 157–183. Springer, New York, NY, USA.
- Stiles, N. R. and Shimojo, S. (2015). Sensory substitution, in: *The Oxford Handbook of Perceptual Organization*, J. Wagemans (Ed.), pp. 775–798. Oxford University Press, Oxford, UK.
- Stiles, N. R., Zheng, Y. and Shimojo, S. (2015). Length and orientation constancy learning in 2-dimensions with auditory sensory substitution: the importance of self-initiated movement, *Front. Psychol.* **6**, 842. DOI:10.3389/fpsyg.2015.00842.
- Striem-Amit, E. and Amedi, A. (2014). Visual cortex extrastriate body-selective area activation in congenitally blind people “seeing” by using sounds, *Curr. Biol.* **6**, 687–692.
- Striem-Amit, E., Guendelman, M. and Amedi, A. (2012). ‘Visual’ acuity of the congenitally blind using visual-to-auditory sensory substitution, *PLoS One* **7**, e33136. DOI:10.1371/journal.pone.0033136.
- Tversky, B. (2003). Structures on mental spaces: how people think about space, *Environ. Behav.* **35**, 66–80.
- Vandenberg, S. G. (1971). *Mental Rotation Test*. University of Colorado, Boulder, CO, USA.
- Wan, C. Y., Wood, A. G., Reutens, D. C. and Wilson, S. J. (2010). Early but not late-blindness leads to enhanced auditory perception, *Neuropsychologia* **48**, 344–348.
- Ward, J. and Meijer, P. (2010). Visual experiences in the blind induced by an auditory sensory substitution device, *Consc. Cogn.* **19**, 492–500.
- West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging, *Psychol. Bull.* **120**, 272–292.
- 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.
- 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.