Sensory Substitution *is* Substitution

JEAN-RÉMY MARTIN AND FRANÇOIS LE CORRE

Abstract: Sensory substitution devices (SSDs) make use of one substituting modality (e.g. touch) to get access to environmental information normally accessed through another modality (e.g. vision). Based on behavioural and neuroimaging data, some authors have claimed that using a vision-substituting device results in visual perception. Reviewing these data, we contend that this claim is untenable. We argue (i) that the kind of information processed by a SSD is metamodal, so that it can be accessed through any sensory modality and (ii) that the phenomenology associated with the use of a SSD is best described in terms of spatial phenomenology, only.

1. Introduction

Sensory substitution devices (SSDs) make use of one sensory modality (substituting modality) in order to get access to a certain type of information normally accessed through another modality (substituted modality). For instance, the *Tactile Visual Sensory Substitution* device (henceforth, TVSS) makes use of a head-mounted video camera capturing environmental information which is transduced into pin vibrations on one part of the body (Bach-y-Rita *et al.*, 1969; Bach-y-Rita, 1972). Another instance is the *Prosthesis Substituting Vision by Audition* (henceforth, PSVA) that translates black-and-white images from a head-mounted video camera into sounds that the subject hears through headphones in real time (Capelle *et al.*, 1998). More precisely, the PSVA uses different sound maps for specific parameters of the scanned object; a single sinusoidal tone is assigned to each pixel of the artificial retina of the camera: frequency codes for the x- and y-axis, i.e. frequencies increase from left to right and from bottom to top, and loudness codes for the grey-scale level of each pixel, i.e. the brighter the image the louder the sounds.[1]

We are very grateful to Malika Auvray, Roberto Casati, Ophélia Deroy, Jérôme Dokic and Elizabeth Pacherie for insightful comments, to the participants of the 16th Annual Conference of the British Postgraduate Philosophical Association, held at the University of Edinburgh, 3–6 September 2012. We are also thankful to the British Society for the Philosophy of Science for having awarded an earlier version of this paper. This work has been supported by the two following grants: ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL.

Address for correspondence: Jean-Rémy Martin, Ecole Normale Supérieure, Département d'Etudes Cognitives, Institut Jean Nicod (CNRS, ENS, EHESS), 29 Rue d'Ulm, F-75005 Paris (FRANCE).
Email: Jean-Remy.Martin@ens.fr

[1] Some SSDs have been designed for other sensory losses, such as bilateral vestibular damage (Tyler *et al.*, 2003), deafness, the loss of tactile sensation (Bach-y-Rita *et al.*, 2003), and the absence of pain feelings (Brand and Yancey, 1993—quoted in Auvray and Myin, 2009). In this article, we shall focus on *visual* substitution only.

Mind & Language, Vol. 30, No. 2 April 2015, pp. 209–233.

After a training period with these devices, SSD-users are able, *inter alia*, to recognize and localize a number of ecologically relevant items such as remote everyday objects (Auvray *et al.*, 2007), to discriminate spatially arranged patterns of horizontal and vertical bars displayed on a monitor screen (Arno *et al.*, 2001a; Poirier *et al.*, 2007a) or letters with different orientations (Sampaio *et al.*, 2001), as well as to discriminate overlapping objects to judge their approximate distance, and, in some cases, to extract their depth information (e.g. Auvray *et al.*, 2005, 2007; Bach-y-Rita *et al.*, 1969; Bach-y-Rita, 2004; Renier *et al.*, 2005a). SSD-users are also able to navigate to some extent in constrained environments and to avoid and walk around obstacles (e.g. Chebat *et al.*, 2011; Collins, 1985).

Along with these abilities, SSD-users report what we will call an 'experiential shift'. After training with, e.g. a TVSS, (some) participants report not feeling the tactile stimulations on their body produced by the SSD anymore. Instead, they report to be aware of the external remote objects themselves. This is referred to as *distal attribution* (Auvray *et al.*, 2005). For instance, Bach-y-Rita *et al.* (1969) observe: 'Our subjects spontaneously report the external localization of stimuli in that sensory information seems to come from in front of the camera, rather than from the vibrotactors on their back' (p. 964).

In the current philosophical and scientific literature, perception with a SSD (or SSD-perception) has been explained in at least two different ways (Auvray and Myin, 2009; Deroy and Auvray, 2012; Deroy and Auvray, 2014): on one side, the 'Deference View' argues that devices substituting for vision, gives rise to *visual* perceptual states (e.g. Bach-y-Rita *et al.*, 1969; Hurley and Noë, 2003; Noë, 2004; O'Regan and Noë, 2001; Renier *et al.*, 2005a, 2005b, 2006); on the other side, the 'Dominance View' argues that such devices give rise to perceptual states belonging to the substituting modality (touch, in the case of the TVSS and audition, in the case of the PSVA).

Here we propose an alternative view we will refer to as the 'Metamodal Spatial View'. Like the Deference and Dominance views we accept that the substituting modality plays a causal role in the production of the experiential shift. However, we will argue that the resulting phenomenology neither belongs to the substituted (Deference) nor to the substituting (Dominance) modality. We propose that SSD-perception gives rise to a kind of spatial phenomenology that is not reducible to either the phenomenology associated with the substituting modality or to the phenomenology associated with the substituted modality.

In the second section, we review some behavioural data that have been thought to be supportive of the Deference view. In particular, we review experiments showing that participants tested with the PSVA are sensitive to illusions, such as the Ponzo illusion, which are usually considered to be visual illusions (Renier *et al.*, 2005b, 2006). We show that such illusions are not, in fact, strictly visual illusions and, therefore, we argue that the sensitivity with a SSD to these illusions does not favour Deference.

In the third section, we review some neuroimaging data apparently favouring Deference (Renier *et al.*, 2005b; Ortiz *et al.*, 2011). These data show that certain visual areas are recruited when participants perform specific tasks with a SSD (e.g. depth perception tasks). We present a view suggesting that (at least some) sensory areas are not modality-specific but computation-specific: they perform a kind of computation (e.g. spatial processing) independently of the modality that conveys the relevant information (Pascual-Leone and Hamilton, 2001). In line with this view, we argue that the recruitment of visual areas during SSD-perception does not show the visual nature of SSD-perception but that SSD-perception illustrates the metamodal nature of these areas considered to be visual.

In the last section, we discuss some phenomenological aspects of SSD-perception. Based on the experiential shift described above, Deference argues as follows: *prima facie* touch does not seem to give access to remote objects, but only to objects one is in contact with, and audition does not seem to give access to fine-grained shape properties—when one hears the sound an object makes, one can recognize and localize the object. However, it is difficult to determine its exact shape by sound information alone. In addition, the combined shape *plus* distance information could be thought of as a kind of information that is accessible only by the visual modality. For Deference, the experiential shift thus leads to *visual* perceptual experiences. For Dominance, the substituting modality simply becomes able, with training, to process information (e.g. distal information) that it seems to be unable to process in normal circumstances (Prinz, 2006).

We shall argue that in the same way that certain sensory areas can be considered as being metamodal, some kind of information can be considered as being metamodal as well. A metamodal information is a kind of information that is potentially accessible by any modality. We will argue, for instance, that the combined shape *plus* distance information is not a type of visual information but a type of metamodal information. Therefore, the fact that participants experientially access such kind of information does not favour Deference. In addition, if one assumes the claim that SSD-users after the experiential shift have distal experiences and do not feel the substituting stimulations anymore, then Dominance seems to be in a delicate position. It is indeed difficult to claim that tactile stimulations, in the case of a visuo-tactile device, are projected in the external space. So, we shall reject Dominance as well, and favour another approach: the 'Metamodal Spatial View'.

2. Does Sensory Substitution Engage Visual Processes?

In order to justify the claim that the use of a SSD gives rise to visual perceptual states, as suggested by the subjective reports of the SSD-users, some advocates of the Deference View argue that visuo-tactile (or visuo-auditory) substitution engages *visual* processes, and not tactile or auditory processes. This claim rests upon two kinds of data: behavioural data concerning the sensitivity of SSD-users to visual illusions and neuroimaging data about the relevant brain areas recruited during SSD-perception. In this section, we focus on the behavioural data.

2.1 SSD-Perception and the Ponzo Illusion

To claim that the behavioural/psychological data are indicative of the visual nature of SSD-perception, the task would have to be such that it is mainly performed by the visual modality in ordinary cases and does not involve conscious learning.

To that purpose, Renier *et al.* (2005b, 2006) tested, with a PSVA, the sensitivity of SSD-users to geometrical illusions such as the Ponzo illusion (2005b) and the Vertical-Horizontal Illusion (2006).[2] In the Ponzo illusion, two horizontal lines are drawn across two converging lines, and the upper line appears longer than the lower line while the two lines are, in fact, of the same length. This illusion could be thought of as being strictly visual, given that it depends on perspective cues (Rock, 1995, ch. 6). For instance, according to one particular account, the Ponzo illusion reflects the involvement of size constancy mechanisms: the converging lines act like perspective cues for the visual system, so that the upper line appears farther away than the lower line; but, given that the two lines have the same length then the visual system interprets the upper line as being longer than the lower one. Consequently, according to Renier *et al.* (2005b), if the user of a PSVA is shown to be sensitive to the Ponzo illusion, which is thought to be related to purely visual mechanisms, then the resulting PSVA-perception of that user could be qualified as visual perception.

The main results of Renier *et al.* (2005b) are the following: (i) the early blind participants were not sensitive to the Ponzo illusion;[3] (ii) the blindfolded sighted participants were sensitive to the Ponzo illusion; (iii) the blindfolded sighted participants were less sensitive than the control group, composed of non-blindfolded sighted participants. The non-sensitivity of the early blind participants to the Ponzo illusion is interpreted by the authors as the fact that early blinds were not sensitive to the perspective cues, given that such sensitivity could depend on previous visual experiences they never had (Gregory and Wallace, 1963; Yonas *et al.*, 1978—quoted in Renier *et al.*, 2005b). Conversely, given that the blindfolded sighted participants already had previous visual experiences, they were sensitive to perspective cues.

However, how did blindfolded sighted participants access the perspective cues with the PSVA? Renier *et al.*'s answer is that the process underlying the sensitivity to the Ponzo illusion *via* a PSVA reflects a *process of visualization* (p. 865).[4] From that explanation, the authors conclude two things: first, that 'mental imagery could be the common cognitive process shared with vision, playing a role in the optical illusions with the PSVA' (p. 865); second, that 'perception induced by a sensory-substitution device shares perceptual processes with vision. These processes

[2] Here, we focus on the results concerning the Ponzo illusion, but the criticisms below apply to both illusions.

[3] Among the nine early blind participants, seven were congenitally blind and two lost vision before their 20th month of life. These participants are here called 'early blind' because of their lack of visual experience.

[4] The fact that blindfolded participants are *less* sensitive than the control group only suggests, according to Renier *et al.* (2005b), that the perceptual experience is just of a poorer quality than the experience of the control group.

can account for the visual nature of perception by sensory-substitution' (p. 866). We now discuss these conclusions.

2.2 Visualization is not Vision

The first conclusion seems to state that in both unmediated normal vision and SSD-perception, the Ponzo illusion involves mental visual imagery because mental imagery is supposed to be the *common* cognitive process. However, in normal vision, the Ponzo illusion seems to be an automatic effect driven both by the stimuli themselves (not by its mental visualization), and by perceptual or cognitive non-mental imagery based processes (e.g. size constancy) (Rock, 1995). In normal vision, mental imagery is likely not necessary for the illusion to be elicited, so that visualization is not common to both SSD-perception and normal vision.[5] The conclusion, drawn from the use of a mental visual imagery strategy, that SSD-perception is visual in nature is thus unguaranteed.

However, by precluding the inference from visualization to the visualness of SSD-perception, we are not saying that visualization is not involved during SSD-perception, or that visualization is not associated with perceptual processes (e.g. size constancy mechanisms). Accordingly, nothing seems to preclude the possibility that the *perceptual* processes involved during SSD-perception are visual—as assumed by Renier *et al.*'s second conclusion—while being independent of a visualization process. Such a possibility should be resisted, as well.

2.3 The Ponzo Illusion: A Multimodal Illusion

Several studies have shown that the Ponzo illusion, as well as other related geometrical illusions (e.g. Müller-Lyer) could not strictly depend on visual mechanisms (Gentaz and Hatwell, 2004). As already suggested by Pasnak and Ahr (1970): 'explanation of illusions [such as the Ponzo illusion] must move away from the eye […] central factors must be importantly involved [in such] illusions rather than properties of the receptor system' (p. 154).

In particular, it has been shown that, in some conditions, the Vertical-Horizontal Illusion and the Ponzo Illusion can be experienced to the same extent in the tactile-haptic modality (e.g. Casla *et al.*, 1999; Gentaz and Hatwell, 2004; Heller *et al.*, 2002; Suzuki and Arashida, 1992). In addition, it also has been shown that they are equally and identically present in early blind people and that they usually depend on the same modulating factors as in vision (Gentaz and Hatwell, 2004). For instance, in the visual version of the Müller-Lyer illusion, the more acute the

[5] In addition, some evidence suggests that the Ponzo illusion cannot be reproduced in mental imagery without previous perceptual exposure with the illusion (see Giusberti *et al.*, 1998). This suggests that a process of visualization is not sufficient to induce an illusion such as the Ponzo illusion.

angles formed by the arrowheads and the segment, the stronger the illusion; and that is true of the tactual-haptic version of the illusion as well (Gentaz and Hatwell, 2004).

Furthermore, Hanley and Goff (1974) have shown that size constancy mechanisms can be at work in the haptic modality when two sticks are used. The use of the sticks make it possible to replicate the visual case in which the visual angles produced by closer or more distant objects are different (the further away an object, the smaller the visual angle). This visual angle information could be used by the system to evaluate the distance of an object and then 'infer' its actual size. In the stick condition, blind and blindfolded sighted participants manifested size constancy showing that the system can use the angle information in another modality than the visual modality. This gives some evidence that size constancy might not be a specifically visual mechanism.

Therefore, as long as the Ponzo illusion or the Vertical-Horizontal illusion are not strictly visual illusions or do not engage strictly visual processes, it is hard to conclude from the fact that some participants are sensitive to these illusions with a PSVA to the visual nature of SSD-perception.

Nevertheless, if it is admitted that the Ponzo illusion is, in fact, a multimodal illusion, and if early blind people are sensitive to these illusions in the haptic-tactile modality, then one should predict the early blind people to be sensitive to the Ponzo illusion with a PSVA. We thus have to explain why early blind people were not sensitive to the Ponzo illusion in Renier *et al.*'s (2005b) study.

2.4 Early Blind People and the Ponzo Illusion

We propose that the non-sensitivity of early blind participants to the Ponzo illusion could result from their lack of experience with 2D-representations (e.g. paintings) of the 3D-structure of the external world—rather than from their lack of visual experience *simpliciter*, as suggested by Renier *et al.* (2005b). More precisely, the Ponzo illusion is supposed to represent a real 3D-scene, where the parallel lines converge in depth (e.g. if you look at straight railway tracks, they look as if they were converging towards each other, while in fact they are parallel). So, in order to represent such depth cues in a 2D-picture, you have to make the lines present in the picture converging towards each other, as do the inducing lines in the Ponzo illusion. Now, congenitally blind people may have a good enough 3D-representation of the external space acquired through audition, proprioception and spatial navigation—and that could explain why they can be sensitive to geometrical illusions when they are presented in non-visual modalities. In contrast, they are largely unfamiliar with external public 2D-pictures mimicking the cues the system makes use of in order to represent the 3D-structure of the external space.

In other words, the Ponzo illusion presented in 2D on a monitor screen was likely difficult to interpret by the early or congenitally blind participants as representing two parallel lines converging in depth. Interestingly, as noticed by Renier *et al.* (2005b), some of the early blind participants 'spontaneously reported that the overall shape of the Ponzo figure was reminiscent of a mental representation of a trestle supporting a table' (p. 865). That suggests that early blind people did not extract the correct 3D-representation from the 2D-depth cues present in the display.

In this section, we have argued that certain suggestive behavioural data in favour of Deference are not conclusive. We now turn to the Neuroimaging data.

3. Neuroimaging Data: What Lesson for SSD-Perception?

The advocates of the Deference View have employed another strategy in arguing that if strictly *visual* areas are found to be activated during SSD-perception, then the resulting perceptual experience is best described as a case of visual experience rather than as a case of tactile or auditory experience. However, in what follows we develop the idea that the areas recruited in SSD-perception could not be strictly visual areas, but metamodal areas instead (Pascual-Leone and Hamilton, 2001). Our claim rests upon data highlighting the metamodal organization of the brain, according to which (at least some) sensory cortical areas are not modality-specific but *computation-specific*, i.e. process a particular kind of information independently of the modality that conveys this kind of information (Pascual-Leone and Hamilton, 2001).

3.1 Background Observations: Crossmodal Activations and the Metamodal Organization of the Brain

Our perceptual experiences of the world are fundamentally multisensory. This potentially implies that our senses substantially interact with each other, rather than being completely separate from one another (at least at a subpersonal level). In fact, crossmodal interactions have largely been experimentally tested, showing for instance that tactile or sound stimuli can influence visual temporal acuity, visual motion perception accuracy and other dimensions of visual perception (for reviews see e.g., Spence, 2011 and Shams and Kim, 2010).

These crossmodal interactions are also reflected at a neuronal level (Stein and Meredith, 1993). Specifically, some brain areas originally considered as unimodal areas can be recruited by other modalities, in both blind and sighted people. In particular, occipital areas can be recruited by auditory or tactile stimuli, manifesting crossmodal activations in sighted people (i.e. the fact that an area normally dedicated to process the stimuli of one modality processes the stimuli of other modalities) and crossmodal plasticity in blind people (i.e. the fact that an area normally dedicated in sighted people to process the stimuli of one modality processes the stimuli of other modalities in blind people) (see, e.g., Amedi *et al.*, 2001; Cohen *et al.*, 1997; Kujala *et al.*, 1995; Lewis *et al.*, 2010; Macaluso *et al.*, 2000; Merabet *et al.*, 2007; Pascual *et al.*, 2005; Sadato *et al.*, 1996; Shams and Kim, 2010).

As an illustration, Macaluso *et al.* (2000) have found, using functional Magnetic Resonance Imaging (fMRI), that the lingual gyrus (a visual area) exhibits increased activity when a tactile stimulation spatially congruent with a visual stimulus is applied on the subject's hand. Even some primary sensory areas, which were thought to be exclusively unimodal, can be crossmodally modulated (Ghazanfar and Schroeder, 2006). For instance, specific tactile tasks, such as the evaluation of say, the orientation of raised-dot patterns can also, under certain conditions, activate the primary visual cortex (Merabet *et al.*, 2007). In blind people, a number of studies have found crossmodal activations of visual areas for (*inter alia*) auditory localization (Collignon *et al.*, 2009), pitch-change discrimination (Kujala *et al.*, 2005), three-dimensional tactile shape recognition (Amedi *et al.*, 2001) and Braille reading (Hamilton and Pascual-Leone, 1998; Ptito *et al.*, 2008).

However, one may wonder if these interactions alter the functional role of those areas. For instance, when a visual area is activated through a non-visual stimulation, does it follow that this area endorses a new function? Interestingly, a number of studies have shown that visual areas keep their specific functional role during crossmodal activations. For instance, the medial temporal area (MT) considered as a visual area preferentially responding to motion stimuli can also be activated by auditory *motion* (simulated by dynamically changing the interaural level difference) and by tactile *motion* both in sighted and blind people (Poirier *et al.*, 2006; Ricciardi *et al.*, 2007).[6] Crossmodal activations could then reflect a real functional engagement of those visual areas in the relevant task, and not just epiphenomenal neuronal activations (Bubic *et al.*, 2010).

In addition, Sathian *et al.* (1997) have found that transcranial magnetic stimulations (henceforth, TMS) applied over the extrastriate visual cortex of sighted participants recruited during the tactile discrimination of gratings orientation significantly reduced their performances (see also Zangaladze *et al.*, 1999 and Sathian, 2005 for a review). TMS applied over the occipital cortex of blind people disrupted their Braille reading abilities (Cohen *et al.*, 1997; Merabet *et al.*, 2004). Collignon *et al.* (2007, 2009) have also shown that TMS applied over the right dorsal extrastriate cortex, which was activated in blind people during sound localization, reduces their performances in that task (and, interferes with the use of a PSVA).

Based on such kinds of results Pascual-Leone and Hamilton (2001) have proposed that the brain could be organized in metamodal *computation*-specific areas rather than in modality-specific sensory areas. That means that (at least certain) cortical areas perform specific types of computations regardless of the modality involved. For instance, some dorsal occipital areas would be involved in processing spatial information, such as localization, irrespective of whether the information is accessed through e.g., the visual modality, the auditory modality or the tactile modality.

Certain kinds of computations may primarily involve brain areas associated with a given modality because they are best accomplished using information typically

[6] These studies have been criticized, however. In particular, crossmodal activations of MT could be an artefact of group averaging. It appears that intra-individual auditory motion responses could have been restricted to an adjacent non-visual area that did not overlap with MT (Lewis *et al.*, 2010).

conveyed through this modality. But suitability is not exclusivity. As Pascual-Leone and Hamilton (2001) claim: '[I]n this view, the "visual cortex" is only "visual" because we have sight and because the assigned computation of the striate cortex is best accomplished using retinal, visual information' (p. 428). The fact that the information acquired through a given sensory channel may be more suitable for certain types of computations may also explain why areas traditionally thought to be unimodal do not always reveal their crossmodal abilities. For instance, the crossmodal potentialities of occipital areas can be silent in a context of visual dominance, but can be unmasked during (even transient) visual deprivation or in special conditions (Merabet *et al.*, 2007; Pascual-Leone and Hamilton, 2001). In fact, after two days of blindfolding, both the striate and the peri-striate cortex of sighted participants were recruited in processing tactile and auditory information (Pascual-Leone and Hamilton, 2001). Accordingly, for Pascual-Leone and Hamilton (2001) crossmodal plasticity, as a result of short- or long-term visual deprivation, may be due to the enhancement of existing crossmodal responses that are normally masked by visual input.

We now propose a metamodal interpretation of the crossmodal activations found during tasks with sensory substitution devices.

3.2 Metamodal Interpretation of Crossmodal Activations during SSD-Sessions

Recent neuroimaging data suggest that certain 'visual' areas are recruited during tasks with SSDs in both blind and sighted participants (e.g. Amedi *et al.*, 2007; Arno *et al.*, 2001b; Bubic *et al.*, 2010; Lacey *et al.*, 2009; Ortiz *et al.*, 2011; Poirier *et al.*, 2006; Poirier, 2007b; Ptito *et al.*, 2005; Renier *et al.*, 2005a; Ricciardi *et al.*, 2007; Tal and Amedi, 2009). For instance, Arno *et al.* (2001b) and Ptito *et al.* (2005) ran PET studies which show that occipital areas in early and congenitally blind participants are recruited by visuo-tactile and visuo-auditory substitution devices, respectively, but not in blindfolded sighted participants, during pattern and orientation discrimination tasks, respectively. An fMRI study conducted by Amedi *et al.* (2007) shows that the lateral occipital tactile-visual area (LOtv), which preferentially responds to visual and haptic shape, is also recruited when sighted and blind participants extract shape information with a visuo-auditory substitution device (in this case, the vOICe)[7]—for further information about that device, see Meijer, 1992).[8] At this point, the Deference/Dominance debate is concerned with the question whether the recruitment of these areas usually considered as visual areas warrants or not the conclusion that the resulting SSD-perceptual states are visual. Renier *et al.* (2005a) and Ortiz *et al.* (2011) respond positively. In contrast, we suggest that the

[7] The vOICe is a visuo-auditory substitution device, which converts visual images into auditory signals (Meijer, 1992). The capital letters 'OIC' stand for 'Oh, I See'.

[8] See also James *et al.*, 2011.

recruitment of such areas does *not* show the visual nature of SSD-perception but rather the metamodal nature of these areas (Amedi *et al.*, 2007; Arno *et al.*, 2001b; Ptito *et al.*, 2005; Ricciardi *et al.*, 2007).

In an fMRI study, Renier *et al.* (2005a) have found significant activations of the fusiform gyrus (bilaterally), the left superior parietal lobule, the precuneus (bilaterally) and the left inferior parietal lobule while participants had to perform depth perception tasks with a PSVA. These areas are considered to be visual areas (Renier *et al.*, 2005a). Consequently, their recruitment by the PSVA (delivering auditory stimuli) suggests at least two possibilities: either these areas are not visual areas but, metamodal areas, or these areas are purely visual areas so that SSD-perception engages visual processes. Renier *et al.* (2005a) opt for the second option given that '[i]n the present study it is obvious that only visual brain areas were recruited during depth perception' (p. 578). In other words, according to Renier *et al.* (2005a) such areas are modality-specific in addition to be computation-specific, leading them to conclude that the recruitment of those areas under SSD-perception '[…] support[s] the concept that perceptions obtained by sensory substitution of vision are visual in nature' (p. 578). However, the idea that the fusiform gyrus, the left superior parietal lobule, the left inferior parietal lobule and the precuneus would be modality-specific is controversial. Certain studies suggest that those areas might actually be multimodal areas rather than modality-specific visual areas: for the fusiform gyrus, see Kung (2007), for the left inferior parietal lobule, see Bushara *et al.* (1999), and for the precuneus, see Cavanna and Trimble (2006).[9]

We now turn to Ortiz *et al.*'s (2011) electroencephalography (EEG) study. Eighteen blind participants were trained during three months (five sessions per week) at discriminating the orientation of a line (horizontal, vertical or oblique) using a tactile piezoelectric device. After the training period, the participants were asked to discriminate pairs of tactile stimuli (e.g. horizontal *versus* oblique).

The post-hoc analysis highlighted a potential correlation between the subjective reports of some participants —'I see phosphenes', 'I see a white line' … (p. 3)— and the activation of occipital areas. More precisely, the authors observe that for all the participants for whom occipital activations occurred these participants also reported, e.g. brightness experiences. Conversely, when no such activations were found, none of the subjective reports mentions brightness experiences at all. Based on these results, the authors conclude that 'activation of the occipital cortex, by tactile stimuli results in visual qualia in some blind subjects' (p. 6). It seems that one may here have some evidence that the recruitment of the occipital cortex by SSD-perception results from the fact that this kind of perception engages visual

[9] Furthermore, visual imagery is also known to activate the same areas as visual perception, including (sometimes) early visual areas (e.g. Kosslyn, 2005). Accordingly, the activations found by Renier *et al.* could actually reflect visual mental imagery rather than truly (visual) perceptual states. Renier and colleagues acknowledge this possibility. However, as already discussed in Section 2.2 we disagree with their view that visualization implies vision.

processes, in that occipital activations correlate with brightness properties, which appears to constitute strictly visual information (i.e. information only accessible by the visual modality).

However, we argue that these crossmodal activations of occipital areas by SSDs might still reflect the metamodal nature of these areas. We are not denying that the recruitment of occipital areas by other modalities than vision (in order to process information requiring the kind of computations achieved by those areas) can give rise to epiphenomenal illusory visual experiences, such as phosphenes. (The triggered phosphenes in Ortiz *et al.*'s 2011 study are illusory in that the stimuli processed by the participants do not instantiate equivalent brightness properties and they are epiphenomenal in that it is far from obvious that they have a causal role in the participants' performances concerning the orientation discrimination task.)

In line with the metamodal hypothesis, we suggest that in Ortiz *et al.*'s (2011) study the occipital areas are recruited in order to achieve the spatial task at stake (namely, orientation discrimination task). As for phosphenes, we suggest that they simply consist in illusory and epiphenomenal visual experiences accompanying the activation of the recruited occipital areas owing to neural 'memories' or traces of associations between spatial and luminance information formed in and kept by those occipital areas from recurrent previous visual experiences (van de Ven and Sack, 2013).

More precisely, occipital areas are largely recruited by vision, which is the most suitable modality for the computation of some spatial information (e.g. orientation) (Pascual-Leone and Hamilton, 2001). Crucially, in vision spatial information is typically accompanied and/or associated with luminance properties (Spence, 2011) and it is likely that the visual modality is the only modality capable to compute such luminance information. Accordingly, for people who have or had (e.g. late blind people) vision, occipital areas are or have been recurrently and essentially activated and recruited by vision and thus by spatial and luminance information. So, we suggest that in Ortiz *et al.*'s (2011) study the activation of occipital areas recruited for the computation of spatial information gave rise, in some people, to illusory brightness experiences because of a retained neural association between spatial and luminance information (van de Ven and Sack, 2013). In other words, the occipital neural computations of the spatial information at stake (delivered here by tactile sensations) triggered phosphenes because usually such occipital neural computations compute or computed (for late blind people) spatial and luminance information together (or, at least, these specific occipital areas recruited for the processing of spatial information are strongly connected with neural assemblies which specifically compute luminance information). To use a metaphor, we might say that tactile stimuli here act like TMS pulses, which, when applied to occipital areas, are known to trigger phosphenes in some conditions (van de Ven and Sack, 2013).

Our proposal finds some corroboration in the fact that, among the participants tested by Ortiz *et al.* (2011) those who have reported having seen phosphenes or luminous lines, either had residual vision ($<3\%$ to $<7\%$ light perception) or were late blind people, which therefore had recurrent visual experiences. In contrast, none of the congenitally blind participants reported having seen phosphenes or luminous lines.[10]

In addition, it has been shown that in some blind participants the stimulation of specific visual areas, which were found to be activated by atypical inputs (i.e. inputs that activate visual areas, but would not in typical cases), produces a type of perceptual experience associated with the atypical inputs, instead of the type of experience usually associated with these cerebral regions. For instance, Ptito *et al.* (2008) found that TMS applied on the occipital cortex of a blind reader induces tactile sensations in their fingers (and not visual sensations). Kupers *et al.* (2006) found similar results in blind people while using a tongue display unit (TDU).[11] In particular, participants reported 'somatopically organized tactile sensations that are referred to the tongue when transcranial magnetic stimulation is applied over the occipital cortex [which is usually associated with visual sensations]' (p. 13256).

In this section we have suggested that the 'visual' areas recruited by SSD-perception may not be, in fact, strictly visual areas, but rather metamodal areas which are involved in the computation of, e.g., spatial information, independently of the modality that conveys this information.

We then suggest that a SSD makes it possible to perform a kind of spatial processing that is normally more suitably performed using visual information, and the areas that are specialized in spatial processing reveal their sensitivity to the substituting stimulations; a sensitivity that was hidden in a context of visual dominance (hence the view that such areas are essentially visual).

At this stage, in order to have a better idea of which kind of phenomenology is associated with SSD-perception, we have to pay close attention to the SSD-users' reports about their own subjective perceptual experience. Indeed, the findings from the TMS studies described above are ambiguous and not completely telling, in that in Kupers *et al.*'s (2006) study for instance, the activation by TMS of the occipital areas of blindfolded sighted participants did not produce tactile sensations, contrary to expectations. Last but not least, the parallel between brain activation and subjective experience should be taken with caution.

4. Is the Phenomenology of SSD-Perception Visual?

So far, we have shown that none of the behavioural/psychological and neuroimaging data support the Deference View. In this section, we claim that the substituting

[10] However, as a whole, Ortiz *et al.*'s (2011) study has to be taken cautiously in that some people among the group that had previous visual experiences or residual vision did not experience visual qualia or occipital activations. So, neither the visual/deferentialist interpretation nor the metamodal spatial interpretation can account for all of their results (or absence of results). An alternative interpretation of these results, which postulate that SSD-perception amount to a synaesthesia phenomenon, will be discussed in Section 4.3

[11] TDU is a visual-to-tactile SSD that uses the tongue (instead of, e.g., the back, or the stomach of the user) as the receptor site of the substituting tactile stimulations (Bach-Y-Rita *et al.*, 1998a, 1998b, 2003).

modality plays a causal rather than a phenomenological role (cf. Section 4.1, fn 12, below) and that the phenomenology of the resulting perceptual experiences should be described in terms of spatial phenomenology, first and foremost. Therefore, we go against Deference but also against Dominance.

4.1 Functional and Experiential Similarities between SSD-perception and Vision

Once more, the general strategy adopted by the advocates of the Deference View goes like this: if it is possible to identify functional and experiential similarities between SSD-perception and vision, then SSD-perception should be qualified as visual. At first glance SSD-perception presents some functional equivalence with unmediated normal vision. As mentioned earlier, SSD-users are able to avoid obstacles, to extract the shape of remote objects, to discriminate overlapping objects, to judge their approximate distances, and, in some cases, to extract depth information, and so on. However, we have suggested that these functional equivalences between SSD-perception and normal vision say nothing about the phenomenology of the resulting perceptual experiences.

Deferentialists are then pushed to claim that SSD-perception gives rise to visual experiences also because SSD-users experientially access what appears to be distinctively visual kinds of information, i.e. information only accessible by the visual modality. Deferentialists think that people access this information *experientially* because after some time of practice with a SSD, an experiential shift happens (section 1). As already mentioned, Bach-y-Rita *et al.* (1969) observe: 'Our subjects spontaneously report the external localization of stimuli in that sensory information seems to come from in front of the camera, rather than from the vibrotactors on their back' (p. 964). Similarly, Hurley and Noë (2003) say: 'objects are reported to be perceived as arrayed at a distance from the body in space and as standing in perceptible spatial relations such as "in front of"' (p. 142). Substituting stimulations are relegated to the margin of consciousness and participants report being 'directly aware' of remote objects detected by the head-mounted camera (Bach-y-Rita *et al.*, 1969; Bach-y-Rita, 2004; Ward and Meijer, 2010). In other words, people attribute the cause of proximal or substituting stimulations to the remote objects processed by the camera and this attribution is accompanied with an experience of objects as being at a distance (experience of distality). On that basis, some authors have suggested that SSD-perception could be qualified as visual. For instance, Bach-y-Rita (2004) states: 'The subjective experience [of blind SSD-users] is comparable (if not qualitatively identical to) vision, *including subjective spatial localization in the three-dimensional world*' (p. 88, our emphasis). The attribution of the substituting stimulations to distal causes has been labelled distal attribution (Auvray *et al.*, 2005), and it is that distal attribution that makes SSD-perception close to vision according to deferentialists, because distal attribution (and the subjective spatial localization of objects in the three-dimensional world that accompanies it) seems to be typical of vision.

More precisely, in ordinary unmediated perception we can experientially access distance information by audition and experientially access shape information by touch, but the combining shape *plus* distance and/or size *plus* distance information seems only accessible by vision. Now, in SSD-perception, it seems that this combination of information becomes accessible for participants. In addition, distal attribution and subjective localization of objects in the three-dimensional world give rise to phenomena that also appear typical of vision such as the phenomenon of occlusion or depth (Bach-y-Rita *et al.*, 1969; Hurley and Noë, 2003, Ward and Meijer, 2010). Occlusion in vision involves opacity, which is a visual notion.

However, we think that the kind of information at stake (shape *plus* distance and size *plus* distance) is not distinctively visual but metamodal, i.e. a kind of information potentially accessible by any modality. In particular, we argue that the combination of shape *plus* distance and/or size *plus* distance information is a kind of metamodal *spatial* information.

The conflation, in the Deference View, between visual and spatial information may come from the fact that for sighted humans this kind of metamodal spatial information is usually processed by vision. And the conflation, in the Dominance View, between tactile (TVSS) or auditory (PSVA) and spatial information may come from the fact that such information is, in the case of SSD-perception, causally accessed through touch or audition. As pointed out by Block (2003) some authors, as Hurley and Noë (2003), 'appear to presuppose that visual phenomenology is shown by the spatial function [of SSDs]. But non-visual senses might be spatial in the same way, e.g. bat sonar' (p. 286). Fish equipped with electroreception can easily access combinations of shape *plus* distance and size *plus* distance information and bats equipped with sonar can easily access combinations of size *plus* distance information (Hara and Zielinski, 2007; see also Hughes, 1999).

Furthermore, having access to occlusion information with a SSD is not distinctively visual. Let us imagine a subject equipped with a visuo-tactile display processing a glass occluding a vertical stick exceeding the width of the glass on both sides. While the subject displaces the camera from left to right, a specific pattern of tactile sensations will be first induced by the non-occluded part of the stick, when the camera arrives at the level of the glass a completely different pattern of tactile stimulations will be induced and, finally, when the camera arrives at the other non-occluded part of the stick, a pattern of tactile stimulations strongly similar to the first non-occluded part of the stick will be induced. If the participants moved the camera continuously, and at constant speed, they will feel, or so we believe, that the glass is actually occluding a single continuing object (i.e. the stick). The extraction of occlusion information here seems to be achieved through the principles of continuity and discontinuity. These principles are not visual (for well-known cases of continuity, e.g. in audition, see Warren, 2008). Finally, depth information too is not a kind of strictly visual information. Once again, electroreception or sonar senses can easily extract depth information (Hara and Zielinski, 2007). There is thus no reason to claim that the information accessed by a visuo-tactile or visuo-auditory substitution device is visual by nature.

Although the metamodal combination of shape *plus* distance or size *plus* distance spatial information does not seem to be given, in everyday-life, by touch or audition, we propose that SSDs provide these modalities with the possibility to process those kinds of information. However, in the next section, we propose that the phenomenology acquired through the processing of metamodal information is not, for instance, tactile (in the case of the TVSS), but purely spatial. As described above, some participants seem to have no feelings of the substituting stimulations anymore while using a SSD. As a consequence, these stimulations might only play a causal role, rather than a phenomenological role, giving rise to purely spatial experiences rather than to tactile or auditory experiences.[12]

4.2 The Phenomenon of Distal Attribution

Distal attribution is a common phenomenon, which is sometimes described as the 'relocation of sensations' (Dennett, 1991; O'Regan and Noë, 2001; Prinz, 2006). More precisely, this relocation of sensations consists in the 'projection' of these sensations at the *source* of stimulations, even when there is a material intermediary between these stimulations and the sensory receptors. This happens in everyday-life. For instance, when you are writing or drawing with a pen, the texture sensations are experienced as being located at the tip of the pen; as if 'your nervous system had sensors out at the tip of the [pen]' (Dennett, 1991, p. 47). The same thing happens for blind people using a stick to navigate in the environment, they feel the texture sensations directly at the tip of the stick (Descartes, 1637/1985). Similarly, when you scrape the floor with your shoes, it is as if the sensations were located under the soles themselves, or when your car drives over some oil spot, you feel the oil spot directly under the wheels (Dennett, 1991; Prinz, 2006).

Experimental studies of the relocation phenomenon include the case of the *Rubber Hand Illusion* (e.g. Botvinick and Cohen, 1998). In this illusion, the experimenter brushes e.g., the left hand of the subject—which is hidden from her view—while synchronously brushing a rubber hand—which is not hidden from her view. After some time, the subject feels the sensations directly on the (seen) rubber hand and not on her (unseen) actual hand. In normal unmediated perception, the sensations are relocated at the source of the stimulations as well. We feel visual sensations at the level of remote objects while the stimulations actually impinge the retina (O'Regan and Noë, 2001). This example can be applied to the case of audition too.

We may think that SSD-perception reflects another case of this phenomenon of the relocation of sensations at the source of stimulations. Now, in the case of the stick used by blind people, we would be very reluctant to deny that the experience remains tactile. So, one can wonder: why should one think that the phenomenal experiences in the case of the TVSS are not tactile as well?

[12] By saying that the substituting sensations have no 'phenomenological role', we mean that these sensations do not lead to a corresponding phenomenology typical of the substituting modality.

The reason is that in the case of the stick or of the rubber hand there is, respectively, a real or apparent connection between the subject's body and these objects. In contrast, in SSD-perception there is no apparent connection. In the former case, the body is 'extended', that is to say the relevant cognitive system considers the objects as part of its body representation (Ehrsson, 2012). This is probably the reason why sensations are relocated, for instance, at the tip of the stick; because the stick is somehow considered as a part (or an extension) of the body (de Vignemont, 2011).

In the case of SSD-perception, the external objects are likely not considered as some part of the body and it seems difficult to make sense of the idea that the proximal tactile sensations delivered by the TVSS are projected in the external space or that the body is extended at the location of the external objects. That is still an open option; but a potentially more plausible alternative is to say, as we do, that the experience of distality acquired after training amounts to a kind of pure spatial experience.

In fact, the experience of distality itself, even in normal unmediated perception, can be thought of as having its own phenomenology. Consider, as an example the case of seeing an object *o* at some distance *x* and hearing the same object *o* at the same distance *x* (e.g. as when you both see and hear your phone). The experience of distality is causally given by vision in the first case and by audition in the second case, but the experience of distality in both cases is the same. However, in such cases it is difficult to isolate the experience of distality from the visual or the auditory qualities that accompany it. In the case of SSD-perception, we are in a special perceptual situation where a shift happens in which the proximal sensations (causally) give rise to an experience of distality which is therefore more salient in comparison to normal perception where the experience of distality is directly given.

4.3 Does SSD-perception allows us to access only Spatial Information?

Is it true that SSD-users can only access metamodal spatial information? In Section 3.2, we described Ortiz *et al.*'s (2011) study that shows that tactile discrimination of orientation information gave rise to brightness experiences (i.e. to phosphenes). Similarly, Ward and Meijer (2010) described some results suggesting that blind SSD-users may access colour information. To the extent that brightness/colour information are intuitively kinds of information that are strictly visual (i.e. only accessible by the visual modality), then one may argue that the metamodal spatial hypothesis is wrong.

First, as we already argued in Section 3.2, the triggered phosphenes in Ortiz *et al.*'s (2011) study are not environmental information accessed by SSD-users, but illusory epiphenomenal experiences resulting from the activation of occipital areas that were recruited for the spatial processing of orientation information. Second, it should be noticed that Ward and Meijer's (2010) survey is based on the subjective reports of only two late blind participants, (PF became blind at 21 years old and CC became blind at 33 years old), and that only one subject (PF) reported having 'seen' colours. Therefore, when PF reports having seen colours, nothing precludes the possibility that she is just inferring or imagining the colour of what she perceives, because of her memory of the colour of the things in question. In fact, PF reported colour experiences only for objects that were familiar to her; as the authors said: 'PF's description of perceiving colours for *known* objects points to the involvement of prior visual knowledge in her phenomenology' (Ward and Meijer, 2010, p. 497, our italics). In other words, it is far from clear that such experiences reflect truly *perceptual* colour experiences triggered in a 'bottom-up way' by stimulations of the substituting modality. It might be that such experiences only reflect top-down mental imagery based on previous visual knowledge (Ward and Meijer, 2010 also discuss this possibility). At least, further experimental data are needed to clarify that point.

Ward and Wright (2012) propose to explain the cases of PF and phosphenes in Ortiz *et al.* (2011) by the hypothesis that SSD-perception would amount to a kind of synaesthesia. In other words, SSD-perception would belong to the substituting as well as to the substituted modality in much the same way that, in synaesthesia, the experience of a sound, say, is associated with the experience of a colour. On the one hand, the synaesthesia and the metamodal spatial hypothesis are not mutually exclusive: one may argue that what the participants can access with a SSD is only metamodal spatial information and that in some rare cases the substituting stimulations trigger illusory experiences belonging to the substituted modality (e.g. phosphenes, see Section 3.2). On the other hand, it is a difficult issue to determine what exactly the nature of the induced experiences in synaesthesia consists in (e.g., the colour experience induced by a sound). That is to say, whether those induced experiences reflect truly *perceptual* experiences or not.[13] In cases of induced colours (e.g. in grapheme-colour synaesthesia) some phenomenological reports and psychophysics studies suggest that such colour experiences are different from the normal non-induced colour experiences (e.g. Edquist *et al.*, 2006; Gheri *et al.*, 2008; Hong and Blake, 2008; Rothen and Meier, 2009). Ward *et al.* (2010) also showed that synaesthesia requires attention contrary to some previous claims (Ramachandran and Hubbard, 2001), and they showed that a group of 9 synesthetes, called projectors, was not any better than a group of 27 synesthetes, called 'associators', in visual search tasks. In a recent fMRI study, Hupé *et al.* (2011) have investigated the specific areas activated by grapheme-colour synaesthesia, and they found 'that none of the individual retinotopic or colour areas responded to synesthetic colours, regardless of the strength of the synesthetic association, and regardless of the way one defines colour [Regions of Interest]' (p. 7). Hupé *et al.*, argue that the synesthetic experience of colours is therefore not equivalent to colour perception.

To conclude this section, note that the Metamodal Spatial View does not require that touch (for instance) should be able to access metamodal information about

[13] Nobody is questioning the reality of synaesthesia as a psychological phenomenon; however, whether the synesthetic experience is fundamentally perceptual or not is a real issue (Hupé *et al.*, 2011).

shape, location and so on with *the same precision* (e.g. with the same resolution) as vision. Sensory modalities (at least, touch, audition and vision) are characterized by a specific sensory bandwidth, 'which refers to the rate at which information from the peripheral sense organs can be transmitted via the afferent pathways to the brain' (Loomis *et al.*, 2012, p. 3). In particular, the spatial bandwidth of vision is much larger than the spatial bandwidth of touch or audition for instance (Loomis *et al.*, 2012). In other words, the latter act like low-pass spatial filter so that the high spatial frequencies that are normally accessible through vision are lost by the use of a visual-to-tactile or visual-to-auditory substitution device. As the authors put it: 'attempting to use some isomorphic spatial mapping from a video camera into the spatial dimensions of touch or hearing inevitably means a huge loss of information' (p. 5). Therefore, it does not imply that a SSD-user will have the exact same contents with SSD-perception than with unassisted visual perception.

5. Concluding Remarks

Some supporters of the sensorimotor theory (Hurley and Noë, 2003; O'Regan and Noë, 2001) argue that SSD-perception is closer to the deferent modality (e.g. vision) than to the substituting modality (e.g. touch). Roughly, the sensorimotor theory considers that perceptual experience occurs when the organism masters laws of sensorimotor contingency. For instance, what it is like to see a pigeon in front of us is constituted by our mastery of some sensorimotor laws expressing the fact that if one moves our eyes, our body or the pigeon itself in a specific way, this will produce specific changes in experience.

The deference hypothesis defended by the sensorimotor theory is based on two observations: first, the post-training phenomenology (i.e. the experience of distality) is more like a visual phenomenology than a tactile or an auditory phenomenology; second, this post-training phenomenology occurs only when the SSD-user herself is allowed to move the camera at will; it does not occur, or the experiential shift does not happen, when it is the experimenter her/himself that moves it (Bach-y-Rita *et al.*, 1969 and Hurley and Noë, 2003). The second point seems to speak in favour of the sensorimotor view of the phenomenology of perceptual experience in that a SSD-user will have a meaningful perceptual experience (e.g. the subject perceives a horizontal line versus random tactile stimulations), only if s/he begins to master sensorimotor contingencies. In addition, the explanation given by the sensorimotor theorist of the alleged visual post-training phenomenology is that the pattern of sensorimotor contingencies mastered in the case of a SSD is similar to those mastered in the case of normal vision. In consequence, the substituting modality in SSD-perception would only have a causal role in enabling the mastering of sensorimotor contingencies that are thought to be *visual* sensorimotor contingencies. If that is true, then the mastering of sensorimotor contingencies would be what shapes the phenomenology of perception.

Nonetheless, we have argued that the phenomenology of SSD-perception is not specifically visual. Therefore, if we are right, the first observation made by the sensorimotor view is erroneous and its argument in favour of the Deference view is undermined at its very basis. That has certain implications for the sensorimotor view in general. First, we can argue that sensorimotor contingencies associated with a SSD are indeed visual and draw the conclusion that sensorimotor contingencies, in general, are therefore not sufficient to explain the phenomenology of perception. Second, we can say that current sensory substitution devices allow for only a partial implementation of visual sensorimotor contingencies. As a result, the phenomenology associated with the experiential shift would be visual but only partially. Finally, we can suggest that there are modality-specific and general sensorimotor contingencies and that sensory substitution devices implement only the second kind.

In conclusion, we have tried to show that neither the behavioural data nor the neuroimaging data support Deference, and that the phenomenological reports of SSD-users could be hardly interpreted in a Dominance or Deference framework. As a result, we proposed that the phenomenology accompanying the experiential shift could be described as a kind of purely spatial phenomenology. Hence, sensory substitution *is* substitution.

Université Paris VI and Institut Jean-Nicod (IEC-ENS-EHESS-CNRS)

References

Amedi, A., Malach, R., Hendler, T., Peled, S. and Zohary, E. 2001: Visuo-haptic object-related activation in the ventral visual pathway. *Nature Neuroscience*, 4, 324–30.

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. *Nature Neuroscience*, 10, 687–9.

Arno, P., Vanlierde, A., Streel, E., Wanet-Defalque, M.C., Sanabria-Bohorquez, S. and Veraart, C. 2001a: Auditory substitution of vision: pattern recognition by the blind. *Applied Cognitive Psychology*, 15, 509–19.

Arno, P., De Volder, A.G., Vanlierde, A., Wanet-Defalque, M.C., Streel, E., Robert, A., Sanabria-Bohorquez, S. and Veraart, C. 2001b: Occipital activation by pattern recognition in the early blind using auditory substitution for vision. *Neuroimage*, 13, 632–45.

Auvray, M., Hanneton, S., Lenay, C. and O'Regan, J. K. 2005: There is something out there: distal attribution in sensory substitution, twenty years later. *Journal of Integrative Neuroscience*, 4, 505–21.

Auvray, M., Hanneton, S. and O'Regan, K. 2007: Learning to perceive with a visuo-auditory substitution system: localisation and object recognition with 'The vOICe'. *Perception*, 36, 416–30.

Auvray, M. and Myin, E. 2009: Perception with compensatory devices: from sensory substitution to sensorimotor extension. *Cognitive Science*, 33, 1036–58.

Bach-y-Rita, P. 1972: *Brian Mechanisms in Sensory Substitution*. New York: Academic Press.

Bach-y-Rita, P. 2004: Tactile sensory substitution studies. *Annals of the New York Academy of Sciences*, 1013, 83–91.

Bach-y-Rita, P., Collins, C.C., Saunders, F.A., White, B. and Scadden, L. 1969: Vision substitution by tactile image projection, *Nature*, 221, 963–64.

Bach-y-Rita, P., Kaczmarek, K. A. and Meier, K. 1998a: The tongue as a man–machine interface: a wireless communication system. *Proceeding of the 1998 International Symposium on Information Theory and its Applications*, 79–81.

Bach-y-Rita, P., Kaczmarek, K. A., Tyler, E.M. and Garcia-Lara, J. 1998b: Form perception with a 49-point electrotactile stimulus array on the tongue: a technical note. *Journal of Rehabilitation Research and Development*, 35(4), 427–30.

Bach-y-Rita, P., Tyler, E.M. and Kaczmarec, A.K. 2003: Seeing with the brain. *International Journal of Human–Computer Interaction*, 15(2), 285–95.

Block, N. 2003: Tactile sensation via spatial perception. *Trends in Cognitive Science*, 7(7), 285–86.

Botvinick, M. and Cohen, J. 1998: Rubber hands 'feel' touch that eyes see. *Nature*, 391, 756.

Bubic, A., Striem-Amit, E. and Amedi, A. 2010: Large-scale brain plasticity following blindness and the use of sensory substitution devices. In M.J. Naumer and J. Kaiser (eds), *Multisensory Object Perception in the Primate Brain*. Berlin: Springer, 351–80.

Bushara, K.O., Weeks, R.A., Ishii, K., Catalan, M-J., Tian, B., Rauschecker, J.P. and Hallett, M. 1999: Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nature Neuroscience*, 2, 8.

Capelle, C., Trullemans, C., Arno, P. and Veraart, C. 1998: A real-time experimental prototype for enhancement of vision rehabilitation using auditory substitution. *IEEE Transactions on Biomedical Engineering*, 45, 1279–93.

Casla, M., Blanco, F. and Travieso, D. 1999: Haptic perception of geometric illusions by persons who are totally congenitally blind. *Journal of Visual Impairment and Blindness*, 93, 583–8.

Cavanna, A.E. and Trimble, M.R. 2006: The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*, 129, 3, 564–83.

Chebat, D.R., Schneider, F.C., Kupers, R. and Ptito, M. 2011: Navigation with a sensory substitution device in congenitally blind individuals. *Neuroreport*, 22, 342–7.

Cohen, L.G., Celnik, P., Pascual-Leone, A., Corwell, B., Falz, L. and Dambrosia, J. 1997: Functional relevance of cross-modal plasticity in blind humans. *Nature*, 389(6647), 180–3.

Collignon, O., Lassonde, M., Lepore, F., Bastien, D. and Veraart, C. 2007: Functional cerebral reorganization for auditory spatial processing and auditory substitution of vision in early blind subjects. *Cerebral Cortex*, 17, 2, 457–65.

Collignon, O., Voss, P., Lassonde M. and Lepore, F. 2009: Cross-modal plasticity for the spatial processing of sounds in visually deprived subjects. *Experimental Brain Research*, 192, 343–58.

Collins, C.C. 1985: On mobility aids for the blind. In D. Warren and E.R. Strelow (eds), *Electronic Spatial Sensing for the Blind: Contributions from Perception, Rehabilitation, and Computer Vision*. Dordrecht: Martinus Nijhoff, 35–64.

Dennett, D. 1991: *Consciousness Explained*. Boston, MA: Little, Brown and Company.

Deroy, O. and Auvray, M. 2012: Reading the world through the skin and ears: a new perspective on sensory substitution. *Frontiers in Psychology*, 3, 457.

Deroy, O. and Auvray, M. 2014: A crossmodal perception on sensory substitution. In D. Stokes, M. Matthen and S. Biggs (eds), *Perception and Its Modalities*, Oxford: Oxford University Press, 327–49.

Descartes, R. 1637/1985: *Philosophical Writings*, J. Cottingham, R. Stoothoff and D. Murdoch, trans. Cambridge: Cambridge University Press.

Edquist, J., Rich, A.N., Brinkman, C. and Mattingley, J.B. 2006: Do synaesthetic colours act as unique features in visual search? *Cortex*, 42, 222–31.

Ehrsson, H.H. 2012: The concept of body ownership and its relation to multisensory integration. In B.E. Stein (ed.), *The New Handbook of Multisensory Processes*. Cambridge, MA: MIT Press.

Gentaz, E. and Hatwell, Y., 2004: Geometrical haptic illusions: the role of exploration in the Müller-Lyer, vertical-horizontal, and Delboeuf illusions. *Psychonomic Bulletin and Review*, 11, 31–40.

Ghazanfar, A. and Schroeder, C.E. 2006: Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, 10, 278–85.

Gheri, C., Chopping, S. and Morgan, M. J. 2008: Synaesthetic colours do not camouflage form in visual search. *Proceedings of the Royal Society B: Biological Sciences*, 275, 841–6.

Giusberti, F., Cornoldi, C., De Beni, R. and Massironi, M. 1998: Perceptual illusions in imagery. *European Psychologist*, 3(4), 281–8.

Gregory, R.L. and Wallace, J.G. 1963: *Recovery from early blindness: a case study*. Quarterly Journal of Psychology. Monograph.

Hamilton, R.H. and Pascual-Leone, A. 1998: Cortical plasticity associated with Braille Learning. *Trends in Cognitive Science*, 2, 168–74.

Hanley, C. and Goff, D. P. 1974: Size constancy in extended haptic space. *Perception and Psychophysics*, 15, 97–100.

Hara, T.J and Zielinski, B.S. (eds) 2007: Sensory systems neuroscience. In A. P. Farrell and C. J. Brauner (eds), *Sensory Systems Neuroscience: Fish Physiology Series*, Volume 25. San Diego, CA: Elsevier Academic Press.

Heller, M.A., Brackett, D.D., Wilson, K., Yoneyama, K., Boyer, A. and Steffen, H. 2002: The haptic Müller-Lyer illusion in sighted and blind people. *Perception*, 31, 1263–74.

Hong, S.W. and Blake, R. 2008: Early visual mechanisms do not contribute to synesthetic color experience. *Vision Res*, 48, 1018–26.

Hughes, H.C. 1999: *Sensory Exotica: A World Beyond Human Experience*. Cambridge, MA: MIT Press.

Hupé, J.M., Bordier, C. and Dojat, M. 2011: The neural bases of grapheme-color synesthesia are not localized in real color-sensitive areas, *Cerebral Cortex*, 22, 1622–33.

Hurley S. and Noë A. 2003: Neural plasticity and consciousness. *Biology and Philosophy*, 18, 131–168.

James, T.W., Stevenson, R.A., Kim, S., VanDerKlok, R.M. and James, K.H. 2011: Shape from sound: evidence for a shape operator in the lateral occipital cortex. *Neuropsychologia*, 49, 1807–15.

Kosslyn, S.M. 2005: Mental images and the brain. *Cognitive Neuropsychology*, 22, 333–47.

Kujala, T., Huotilainen, M., Snikkonen, J., Ahonen, A.I., Alho, K., Hämäläinen, M.S., Ilmoniemi, R.J., Knuutila, J.E.T., Lavikainen, J., Salonen, O., Simola, J., Standertskjöld-Nordenstam, C.G. and Näätänen, R. 1995: Visual cortex activation in blind humans during sound discrimination. *Neuroscience Letters*, 183, 143–6.

Kujala, T., Palva, M.J., Salonen, O., Alku, P., Huotilainen, M., Jarvinen, A. and Naatanen, R. 2005: The role of blind humans' visual cortex in auditory change detection. *Neuroscience Letters*, 379, 127–31.

Kung, C-C. 2007: Expanding the FFA debate: a reevaluation of recent studies and evidence for the multimodal and dynamic fusiform gyrus. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 67(8-B), 4733.

Kupers, R., Fumal, A., de Noordhout, A.M., Gjedde, A., Schoenen, J. and Ptito, M. 2006: Transcranial magnetic stimulation of the visual cortex induces somatotopically organized qualia in blind subjects. *Proceedings of the National Academy of Science*, 103, 35, 13256–60.

Lacey, S., Tal, N., Amedi, A. and Sathian, K. 2009: A putative model of multisensory object representation. *Brain Topography*, 21, 269–74.

Lewis, L.B., Melissa, Saenz, M. and Fine I. 2010: Mechanisms of cross-modal plasticity in early-blind subjects. *Journal of Neurophysiology*, 104, 2995–3008.

Loomis, J., Klatzky, R.L. and Giudice, N.A. 2012: Sensory substitution of vision: importance of perceptual and cognitive processing. In R. Manduchi and S. Kurniawan (eds) *Assistive Technology for Blindness and Low Vision*. Boca-Raton, FL: Taylor and Francis/CRC Press.

Macaluso, E., Frith, C.D. and Driver, J. 2000: Modulation of human visual cortex by crossmodal spatial attention. *Science*, 289, 1206–8.

Meijer, P.B. 1992: An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering*, 39, 2, 112–21.

Merabet, L.B., Swisher, J.D., McMains, S.A., Halko, M.A., Amedi, A., Pascual-Leone, A. and Somers, D.C. 2007: Combined activation and deactivations of visual cortex during tactile sensory processing. *Journal of Neurophysiology*, 97, 1633–41.

Merabet, L., Thut, G., Murray, B., Andrews, J., Hsiao, S. and Pascual-Leone, A. 2004: Feeling by sight or seeing by touch? *Neuron*, 42, 173–9.

Noë, A. 2004: *Action in Perception*. Cambridge, MA: MIT Press.

O'Regan, J.K. and Noë, A. 2001: A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24, 5, 939–1031.

Ortiz, T., Poch, J. and Santos, J.M. 2011: Recruitment of occipital cortex during sensory substitution training linked to subjective experience of seeing in people with blindness. *PLoS One*, 6, e23264.

Pascual-Leone, A. and Hamilton, R. 2001: The metamodal organization of the brain. *Progress in Brain Research*, 134, 427–45.

Pascual-Leone, A., Amedi, A., Fregni, F. and Merabet, L.B. 2005: The plastic human brain cortex. *Annual Review of Neuroscience*, 28, 377–401.

Pasnak, R. and Ahr, P. 1970: Tactual Poggendorff Illusion in blindfolded subjects. *Perceptual and Motor Skills*, 31, 151–4.

Poirier, C., Collignon, O., Scheiber, C., Renier, L., Vanlierde, A., Tranduy, D., Veraart, C. and De Volder, A.G. 2006: Auditory motion perception activates visual motion areas in early blind subjects. *Neuroimage* 31, 279–85.

Poirier C., De Volder A., Tranduy D. and Scheiber C. 2007a: Pattern recognition using a device substituting audition for vision in blindfolded sighted subjects. *Neuropsychologia*, 45(5), 1108–21.

Poirier, C., De Volder, A. and Scheiber, C. 2007b: What neuroimaging tells us about sensory substitution. *Neuroscience and Biobehavioral Reviews*, 31, 1064–70.

Prinz, J. 2006: Putting the brakes on enactive perception. *Psyche*, 12, 1–19.

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(3), 606–14.

Ptito, M., Fumal, A., Martens de Noordhout, A., Schoenen, J., Gjedde, A. and Kupers, R. 2008: TMS of the occipital cortex induces tactile sensations in the fingers of blind Braille readers. *Experimental Brain Research*, 184, 193–200.

Ramachandran, V.S. and Hubbard, E.M. 2001: Synaesthesia—a window into perception, language and thought. *Journal of Conscious Studies*, 8, 3–34.

Renier, L., Collignon, O., Poirier, C., Tranduy, D., Vanlierde, A., Bol, A., Veraart, C. and De Volder, A. 2005a: Cross-modal activation of visual cortex during depth perception using auditory substitution of vision. *Neuroimage*, 26, 573–80.

Renier, L., Laloyauxô, C., Collignon, O., Tranduy, D., Vanlierde, A., Bruyer, R. and De Volder A. 2005b: The Ponzo illusion with auditory substitution of vision in sighted and early-blind subjects. *Perception*, 34, 857–67.

Renier, L., Bruyer, R. and De Volder, A.G. 2006: Vertical-horizontal illusion present for sighted but not early blind humans using auditory substitution of vision. *Perception and Psychophysics*, 68(4), 535–42.

Ricciardi, E., Vanello, N., Sani, L., Gentili, C., Scilingo, E.P., Landini, L., Guazzelli, M., Bicchi, A., Haxby, J.V. and Pietrini, P. 2007: The effect of visual experience on the development of functional architecture in hMT+. *Cerebral Cortex*, 17, 2933–9.

Rock, I. 1995: Perception. *Scientific American Books series*. New York: W.H. Freeman and Co.

Rothen, N. and Meier, B. 2009: Do synesthetes have a general advantage in visual search and episodic memory? A case for group studies. *PLoS One*, 4:e5037.

Sadato, N., Pascual-Leone, A., Grafman, J., Ibanez, V., Deiber, M.P., Dold, G. and Hallett, M. 1996: Activation of the primary visual cortex by Braille reading in blind subjects. *Nature*, 380, 526–8.

Sampaio, E., Maris, S. and Bach-y-Rita, P. 2001: Brain plasticity: 'visual' acuity of blind persons via the tongue. *Brain Research*, 908, 204–7.

Sathian, K. 2005: Journeying beyond classical somatosensory cortex. *Canadian Journal of Experimental Psychology*, 61(3), 254–64.

Sathian, K., Zangaladze, A., Hoffman, J.M. and Grafton, S.T. 1997: Feeling with the mind's eye. *NeuroReport*, 8, 3877–81.

Shams, L. and Kim, R. 2010: Crossmodal influences on visual perception. *Physics of Life Reviews*, 7(3), 269–84.

Spence, C. 2011: Crossmodal correspondences: a tutorial review. *Attention, Perception & Psychophysics*, 73, 971–95.

Stein, B.E. and Meredith, M.A. 1993: *The Merging of the Senses*. Cambridge, MA: MIT Press.

Suzuki, K. and Arashida, R. 1992: Geometrical haptic illusions revisited: haptic illusions compared with visual illusions. *Perception & Psychophysics*, 52(3), 329–35.

Tal, N. and Amedi, A. 2009: Multisensory visual-tactile object related network in humans: insights gained using a novel crossmodal adaptation approach. *Experimental Brain Research*, 198, 165–82.

Tyler, M., Danilov, Y. and Bach-y-Rita, P. 2003: Closing an open-loop control system: vestibular substitution through the tongue. *Journal of Integrative Neuroscience*, 2, 159–64.

van de Ven, V. and Sack, A. T. 2013: Transcranial magnetic stimulation of visual cortex in memory: cortical state, interference and reactivation of visual content in memory. *Behavioral Brain Research*, 236, 67–77.

de Vignemont, F. 2011: Embodiment, ownership and disownership. *Consciousness and Cognition*, 20, 82–93.

Ward, J., Jonas, C., Dienes, Z. and Seth, A. 2010: Grapheme-colour synaesthesia improves detection of embedded shapes, but without pre-attentive 'pop-out' of synaesthetic colour. *Proceedings of the Royal Society B: Biological Sciences*, 277, 1021–6.

Ward, J. and Meijer, P. 2010: Visual experiences in the blind induced by an auditory sensory substitution device. *Consciousness and Cognition*, 19, 492–500.

Ward, J. and Wright, T. 2012: *Sensory substitution as an artificially acquired synaesthesia*. Neuroscience and Biobehavioral Reviews.

Warren, R.M. 2008: *Auditory Perception: An Analysis and Synthesis*. New York: Cambridge University Press.

Yonas, A., Cleaves, W. and Pettersen, L. 1978: Development of sensitivity to pictorial depth. *Science*, 200, 77–9.

Zangaladze, A., Epstein, C. M., Grafton, S. T. and Sathian, K. 1999: Involvement of visual cortex in tactile discrimination of orientation. *Nature*, 401, 587–90.