Sensory Substitution and Augmentation Workshop Report

By

Kevin Connolly, Diana Acosta Navas, Umut Baysan,

Janiv Paulsberg, and David Suarez

This report highlights and explores five questions that arose from the workshop on sensory substitution and augmentation at the British Academy, March 26th through 28th, 2013.

1. Does sensory substitution generate perceptual or cognitive states?

Sensory substitution devices (SSDs) deliver information about the environment normally perceived through stimulation in one sensory modality (the "substituted modality"), through the production of stimulation in another sensory modality (the "substituting modality"). Most SSDs aim to substitute for vision, and consist of a video camera that feeds information into a conversion unit which then converts that information into auditory or tactile stimuli. Such devices are often used by the blind to assist in their autonomous navigation of the world. For instance, tactile-vision sensory substitution (TVSS) devices are SSDs that convert patterns of luminance picked up by a camera into isomorphically-organized tactile stimuli which are delivered to the skin through a matrix of solenoids usually mounted on the back, or a matrix of electrodes held on the tongue.

There is no doubt that sensory substitution devices can convey information to subjects using them. But is the information conveyed to subjects via SSDs perceptual information? Following Ophelia Deroy and Malika Auvray, we may call the assumption implicit in current discourse that the use of SSDs is akin to the appropriation of a sensory modality, the *perceptual assumption*. As Deroy and Auvray put the perceptual assumption:

> [T]he perceptual assumption considers that sensory substitution follows what occurs with canonical cases of perception through one of the typical sensory modalities, that is as specialized channels for transducing external information. As spelled out by Grice (1962), perceiving through each of these specialized sensory routes typically starts out with specific kinds of receptors being stimulated by certain kinds of stimuli; the information is then further processed (at least at an early stage) by dedicated sensory mechanisms that finally deliver a representation of a certain kind of object or properties or leads to specific responses.

One way to settle whether the perceptual assumption is true is to ask whether the information conveyed by SSDs is the output of a perceptual mechanism or not. Deroy and Auvray argue that since the use of SSDs fails to meet the conditions required for the constitution of an appropriation of a perceptual mechanism (stimuli, receptors, processes, and outputs), the perceptual assumption is false. Deroy and Auvray propose an alternative model for understanding the use of SSDs, according to which the use of SSDs is a *cognitive* extension of existing perceptual skills such as reading.

In her reply to Deroy and Auvray, Jennifer Corns detailed three different ways in which we can read the perceptual assumption: (i) *The strong reading*: The use of an SSD is akin to the appropriation of a particular natural sensory modality, such as vision, or audition; (ii) *The moderate reading*: The use of an SSD constitutes a novel and unique sensory modality akin to the natural modalities. (iii) *The weak reading*: The process involved in the use of an SSD constitutes a perceptual process.

Here, Corns notes that the weaker the claim, the stronger are the demands to reject it. To reject the strong reading we only need to demonstrate that there is a difference between the use of an SSD and the target natural sense modality. To reject the moderate reading we need to demonstrate that there is a difference between the use of SSDs and natural sense modalities as such. Finally, to reject the weak reading, one must demonstrate that the information processing involved in the use of an SSD is different than a perceptual process as such.

So one question that arises is whether the evidence cited by Deroy and Auvray is sufficient to reject the perceptual assumption on all three readings. But more importantly, Corns asks the following question: suppose that SSD use differs from perception (in any of the above readings); why assume that the only route is to reject the perceptual assumption? Instead, Corns suggests that in light of the evidence, one might opt to revise one's conception of either: 1) the particular sense modality in question (for example, vision), 2) one's conception of a sense modality, or 3) one's conception of a perceptual process.

At the start of this section we asked whether the information conveyed to subjects via SSDs is perceptual information or not. It has been suggested that the answer to this question depends on whether that information is the output of a perceptual mechanism. In trying to provide an answer to the second question, the discussion revolved about the perceptual assumption with the thought that if the perceptual assumption is false then the information conveyed via SSD is not the output of a perceptual mechanism. But providing an answer to our second question may not be as straightforward as one would have hoped. At the heart of the problem is the following question: do the data about the use of SSDs warrant a rejection of the perceptual assumption (on any of its readings) or a revision of our conception of perception?

References:

Deroy, O., & Auvray, M. (2012). Reading the world through the skin and ears: a new perspective on sensory substitution. *Frontiers in Psychology*, *3*(457), 1–13. doi:10.3389/fpsyg.2012.00457

Grice, H. P. (1962). "Some remarks about the senses," in Analytical Philosophy (First Series), ed. R. J. Butler (Oxford: Basil Blackwell), 248–268.

2. What can sensory substitution tell us about perceptual learning?

Eleanor Gibson defines perceptual learning as "any relatively permanent and consistent change in the perception of a stimulus array, following practice or experience with this array" (1963, p. 29). Are cases of sensory substitution cases of perceptual learning? The answer to this question depends on whether the incorporation of a sensory substitution device yields a perceptual change, rather than a strictly cognitive change. Suppose we assume a perceptual change—one that is permanent, consistent, and results from practice. What would sensory substitution tell us, then, about perceptual learning?

Fiona Macpherson suggested one way in which we might go about answering this question. We might proceed by comparing and contrasting the learning that goes on in sensory substitution devices with other kinds of perceptual learning. For instance, we might compare sensory substitution learning with the learning that occurs when a blind person with cataracts has an operation to get them removed. Several people took up this comparison throughout the conference.

Citing Pawan Sinha's work, Amir Amedi argued that the blind can develop crossmodal matching very quickly (although they do not have it immediately upon having their cataracts removed), and that the same is true with the vOICe.[1] On the other hand, he continued, even several months after surgery, Sinha's subjects fail at visual parsing—the ability to tell whether something is one or more objects. In contrast, in just seventy hours of training with the vOICe, subjects are able to do visual parsing. Mohan Matthen later added another point of disanalogy between sensory substitution cases and cataract removal cases. In cataract removal cases, there is no problem with locating the stimulus. In contrast, with sensory substitution devices, there is the

[1] As Kiverstein, Farina, and Clark describe it, "The vOICe is a visual-auditory substitution device that works by transforming images from a digital camera embedded in a pair of sunglasses into auditory frequencies ("soundscapes"), which the user hears through headphones" (forthcoming).

additional step of distal localization, that is, locating the distal stimulus based on the proximal stimulus that one experiences. As Malika Auvray and Ophelia Deroy have shown, distal localization requires training. So unlike in the cataract removal cases, with sensory substitution devices, localization is not automatic.

The comparison between sensory substitution cases and cataracts removal cases is just one way in which we can find out about perceptual learning. Macpherson suggested three additional ways. First, we might compare and contrast different sensory substitution devices, as well as thinking about what happens when you use them in tandem. The former is important because some sensory substitution devices offer unique instances of perceptual learning. For example, Peter Konig argued that the Feelspace belt enlarges egocentric space about ten fold, allowing subjects to interact with objects well beyond their visual range. Other sensory substitution devices do not do this. Second, we might compare and contrast sensory substitution devices with inverting lenses. For instance, how is substituting a modality different from inverting the information in an existing one? Third, we might think about very basic sensory substitution systems, and ask if the kind of learning that goes on in those cases is the same as the kind of learning that goes on in more complex sensory substitution cases. For instance, in his commentary on Amir Amedi's talk, Derek Brown spoke of Pavlov's bells as being a very basic form of sensory substitution. We might also think about braille and sign language as kinds of sensory substitution devices. Again, we can ask in each of these cases whether ask if the kind of learning that goes on in it is the same as the kind of learning that goes on in sensory substitution.

What then can comparing and contrasting the learning that goes on in sensory substitution devices with other kinds of perceptual learning teach us about perceptual learning? Perhaps the best way to answer this is not by trying to come up with a single, unified thing that such devices teach us. Rather, we might treat each device on its own as an instance of perceptual learning, with perhaps something unique to teach us about what is possible in perceptual learning. We can then be sensitive to differences in kinds of sensory substitution devices, the different training regimes involved, and the different subjects who are using them (whether they be early blind or late blind, for instance). This allows us to abide by Fiona Macpherson's advice that when thinking about sensory substitution devices, we should pay close attention to the fine-grained details of the cases.

References:

Gibson, E. J. (1963) "Perceptual learning." *Annu. Rev. Psychol.* 14, 29–56.

Kiverstein, J., Farina, M., and Clark, A. (forthcoming). "Substituting the Senses." In Mohan Matthen (ed.), *The Oxford Handbook of the Philosophy of Perception*. Oxford University Press.

3. How does sensory substitution interact with the brain's architecture?

One of the important questions about the working of sensory substitution devices (SSDs) is how these devices feed information to the brain. The way in which the brain processes information delivered by these devices could potentially give us clues as to the nature of the experience that is generated by these devices. There are two interrelated questions that are relevant to this topic. First, how do SSDs interact with the pre-existing brain architecture? And second, does this interaction reveal anything about the experiences of SSD users?

On the first question, we might start by asking whether sensory substitution devices are capable of feeding information to the brain systems that normally process visual information. The answer is that, not only do SSDs feed information to the visual areas of the brain, but (as

Maurice Ptito mentioned in his talk) they also make use of the parallel architecture of the visual system. That is, they feed information to both the ventral and the dorsal streams, following a pattern of task-dependent specialization. This shows that the continued use of SSDs exploits the plasticity of the brain, in a way such that, after training, the brain is capable of processing this new kind of information making use its pre-existing systems and modules. During periods of training there is a significant alteration in the subjects' sleeping behavior, which is normally indicative of learning processes taking place via neuronal plasticity. After such periods, the brain activity caused by the use of the devices presents interesting patterns.

In the workshop, Peter Konig presented on the Feelspace belt (which is designed to help subjects orient in space by indicating the magnetic North through vibrations around subjects' waist), He reported that after a training period of six weeks, the belt causes activation in both the right supplementary motor cortex and some parietal areas that are typically involved in processing egocentric spatial information for navigation. This is accompanied by an alteration in subjects' phenomenal experience of space. Subjects' reports show that the belt has a positive influence on their subjective experience of spatial navigation: trained users of the belt reported an improved orientation and navigational ability in unknown territories, along with a continuous improvement of their knowledge of spatial relations. However, these reports contrast with the behavioral data which showed only a minor improvement in tasks of navigation and orientation.

Laurent Renier reported on the PSVA, a visual-auditory SSD similar to the vOICe. He presented evidence that after a training period, when sighted subjects use the PSVA to perform visual tasks of object location, distance estimation and object recognition, there was activation of the visual areas that are normally recruited for the performance of this tasks by use of visual information. At the behavioral level, it was shown that early and congenitally blind subjects (unlike late blind and sighted subjects) are insusceptible to visual illusions induced by effects of perspective. This indicates that the perception of perspective is dependent on strictly visual perception, and cannot be acquired. These behavioral studies show that even though the brain is capable of rewiring in order to use the SSDs, there seems to be a limitation of the way in which SSDs can exploit bran plasticity: Some processes, like the perception of perspective, are dependent on a previous training of the visual system.

Maurice Ptito presented on the tongue display unit (TDU)—a visual-tactile SSD that stimulates the tongue. He explained that when trained blind subjects use a TDU for visual tasks, like shape, motion and orientation identification, the normal visual areas are recruited. His results show that when blind subjects use the TDU to perform visual tasks that are known to activate the dorsal and ventral visual streams in the sighted, they activate the same brain areas. This suggests that motion and shape processing are organized in a supramodal manner in the human brain, and that vision is not necessary for the development of the functional architecture characteristic of motion and shape processing areas.

Ptito also used the TDU in spatial navigation tasks and showed that in contrast to blindfolded sighted subjects, blind subjects activated their visual cortex and right parahippocampus during navigation, suggesting that in the absence of vision, cross-modal plasticity permits the recruitment of the same cortical network used for spatial navigation tasks in sighted subjects. This can be explained by saying that training with the TDU caused the brain to rewire in a way that makes it capable to process tactile stimuli as though they were visual. However, it is known that when a sensory modality is damaged, its part of the sensory cortex normally adapts to represent other kinds of sensory stimuli. So it is not clear whether the activation of visual cortex is a product of the training with the TDU or if it was there before, and the TDU merely exploits this preexisting rewiring of the visual areas.

Amir Amedi's talk made clear an idea that was suggested by other talks: the activation of visual areas is not stimulus-dependent, but task-dependent. Fixing the stimuli but changing tasks resulted in different areas being activated. On the other hand, his results indicate that training increases the selective recruitment of specialized visual areas for certain tasks. That is, there is an increasing specialization of visual areas for the performance of certain tasks, which is driven by training with the SSDs. So, it seems that learning to use SSDs causes a rewiring of the sensory cortex, which allows subjects to perform visual tasks, by processing other kinds of stimuli (tactile and auditory). It is not clear up to what point this rewiring is caused by the training, and how much builds on a preexisting adaptation of the visual areas for the processing of other stimuli.

Finally, there is the question whether this interaction between the brain and SSDs reveals any clues about the kind of experience that is generated by the devices. Even though there seems to be a correlation between the activation of visual areas for perceptual tasks and reports of changes in the subjects' experience, most researchers are skeptical of that claim that this means the subject is having a visual experience. At the conference, Kevin O'Regan claimed that changes in the subjects' experience should not be explained in terms of cortical plasticity, but rather in terms of the acquisition of sensorimotor skills and knowledge. Fiona Macpherson also rejected this approach, claiming that since visual areas also serve non-visual functions, the activation of these areas doesn't settle whether users of SSDs have visual experiences or not.

4. Can normal non-sensory feelings be generated through sensory substitution?

Perception provides us with sensory contents like shapes, colors, flavors, and smells. In ordinary perception, these are accompanied by a variety of non-sensory feelings as well. Perception arouses *noetic feelings* concerning the presence or familiarity of what one is perceiving, *emotional affects* such as love and hate, *hedonic responses* such as pleasure and pain, and *aesthetic feelings* of beauty and ugliness. As Jerome Dokic noted in his presentation, although some non-sensory feelings seem to be naturally associated with specific sensory contents, non-sensory feelings do not seem to be supervenient on the sensory contents they accompany, since non-sensory feelings can vary while sensory contents remain the same.

It has been reported that the use of sensory substitution devices (SSDs) fails to elicit some of the non-sensory feelings expected in normal cases of perception. Though many users of SSDs are eventually able to attribute stimuli in substituting modalities to distal objects, it takes training—and, in many cases, explicit instruction in the mapping between substituting and substituted modalities—to generate the noetic feeling that substituted stimuli are due to the presence of real objects. In addition, Bach-y-Rita et al. report an "absence of qualia" (i.e., absence of affective responses) in trained, adult users of tactile-visual sensory substitution systems (TVSS):

> [W]e found that while experienced blind TVSS subjects could perceive faces and printed images, they were very disappointed when perception was not accompanied by qualia: A Playboy centerfold carried no emotional message, and the face of a girl-friend or a wife created an unpleasant response since it did not convey an affective message. (2003, p. 293; see also, Bach-y-Rita, 2002)

Deroy and Auvray (2012), citing Lenay et al. (2003), report that hedonic responses also fail to transfer from substituted modalities: "[S]hapes perceived in one sensory modality are not directly associated to pleasures or pains felt while perceiving the same shape in another sensory modality" (p. 4). Finally, it seems likely that SSDs will face difficulties in transferring aesthetic responses between modalities. Mohan Matthen raised the question of whether visual-to-auditory sensory substitution could preserve the aesthetic properties of visual stimuli. Would a visual-to-auditory SSD allow us to *hear* as beautiful, what is normally *seen* as beautiful? In response, Malika Auvray speculated that even though such a transfer of responses has not yet been found, it is possible that emotional responses to stimuli could be generated through further training.

This suggestion follows Bach-y-Rita's (2003; 2002) hypothesis that SSDs fail to generate normal non-sensory feelings because the developmental and learning processes which normally result in the association of non-sensory feelings with sensory contents has not taken place. Moreover, he suggests that such associations could be developed through further training and experience in the use of SSDs. Bach-y-Rita suggests that lack of normal affective responses in SSD-use,

> may be compared to the acquisition of a second language as an adult. The emotional aspects of the new language are often lacking, especially with emotionally charged words and expressions, such as curse words. It appears that both spoken language and other sensory messages require long experience within the context of other aspects of cultural and emotional development to be able to contain qualia. (Bach-y-Rita, 2002, p. 510)

In support of this, Bach-y-Rita reports that "Systems for blind babies … have already provided some suggestive evidence for the development of qualia, such as the infant's smile upon perceiving the mother's approach" (Bach-y-Rita, 2002, p. 510). And, as Deroy and Auvray note, "there are similar reports of the absence of emotion and meaning felt by persons blind from birth who recover sight following the removal of cataracts" (2012, p. 4).

But is insufficiency of training or experience really the only stumbling block for generating normal non-sensory responses? In his commentary on Jonathan Cohen's presentation, Charles Spence pointed out that flavor and olfactory substitution devices may lack a point if they cannot duplicate the normal hedonic responses to tasty or fragrant stimuli. The problem is that it seems plausible to think that delivering the relevant information (through, say, vision, audition, or touch) would not generate the pleasure we associate with the flavor of chocolate or the smell of a rose. Moreover, it seems unlikely that further training with such a device *could* succeed in associating visual, auditory, or tactile representations of “tasty” or “fragrant” stimuli with a normal hedonic response. All of this suggests that merely receiving the information normally provided by a sensory modality may not be sufficient for the duplication of that modality’s affective/hedonic components. And, in the case of flavor and olfaction, in particular, the hedonic component seems quite central to the normal functioning of the modality, such that sensory substitution would lose much of its point if the normal hedonic responses could not be duplicated in the substituting modality.

This raises the question of how to conceive of the relationship between affective/hedonic responses and sensory modalities. Are they essential to the phenomenal feel of a sensory modality? In the absence of such responses, would we still have the same modality? What one says in response to these questions may bear on the truth of representationalism. Suppose one holds the *representationalist* view that sensory modalities are to be distinguished by the information they carry about the external world. It would follow that all that is needed for the substitution of sensory modality is the right kind of informational input stream. But if duplicating the informational stream doesn’t duplicate the affective/hedonic responses that normally accompany the functioning of a sense, then have we really substituted for that sense?

References:

Bach-y-Rita, P. (2002). Sensory substitution and qualia. In A. Noë & E. Thompson (Eds.), *Vision and mind: selected readings in the philosophy of perception* (pp. 497–514). Cambridge, MA: MIT Press.

Bach-y-Rita, P., Tyler, M. E., & Kaczmarek, K. A. (2003). Seeing with the brain. *International Journal of Human-Computer Interaction*, *15*(2), 285–295.

Deroy, O., & Auvray, M. (2012). Reading the world through the skin and ears: a new perspective on sensory substitution. *Frontiers in Psychology*, *3*(457), 1–13. doi:10.3389/fpsyg.2012.00457

Lenay, C., Gapenne, O., Hanneton, S., Marque, C., & Genouelle, C. (2003). Sensory substitution: limits and perspectives. In Y. Hatwell, A. Streri, & E. Gentaz (Eds.), *Touching for Knowing* (pp. 275–292). Amsterdam: John Benjamins.

5. What are the limitations of sensory substitution?

Many speakers and commentators in the conference mentioned the limitations of sensory substitution, emphasizing in particular that sensory substitution is not the *substitution* of an entire sensory modality, but rather the replication of several features of that modality. Since sensory modalities carry information about a range of properties, an important feature that sensory substitution attempts to replicate is the feature of *conveying information*. But questions arise: About what sorts of properties can and *cannot* sensory substitution convey information? If information about some properties cannot be conveyed by sensory substitution, then does this mean that sensory substitution has serious limitations? Can the information that is conveyed be rich enough to replicate other features of sensory modalities?

As several speakers at the conference pointed out (most notably Charles Spence), the focus of the sensory substitution research has been vision. Vision allows us to gather rich information about our environments. Since the early days of sensory substitution research, sensory substitution devices have been able to convey a fair amount of information about environment. Subjects using sensory substitution devices are reported to recognize objects, point to objects accurately, judge the distances and the sizes of objects, and even make complex pattern discriminations. Based on these, it might be suggested that sensory substitution devices can carry information about, at least, the *common sensible:* namely, motion, shape and size properties of the objects in the environment. However, there seems to be some limitations even in such cases. As Laurent Renier discussed, some sorts of experiences that are related to depth and distance perception cannot be generated with sensory substitution devices in congenitally blind subjects. If these show that depth perception can only be replicated in late-blind and sighted subjects, then there is a problem that sensory substitution researchers should resolve.

Even if information about common sensibles were to be properly conveyed by these devices, there would still be a limitation with respect to the substitution of the experience of *proper sensibles* such as the color, smell and taste properties of objects. Whether this limitation results from the impossibility of replicating the properly perceptual aspects of proper sensibles, or from technical or design-related problems is a question that remains to be settled. Here, a lot hinges on what theories about perceptual experience are true. If one believes, as Jonathan Cohen argued in his talk, that there are good reasons to think that some features of visual experience are emergent, and so do not supervene on the information that is conveyed by sensory substitution devices, then one might think that sensory substitution cannot restore those emergent features simply by delivering the right information. Moreover, this or a similar reason might also explain why there are no well-known examples of the substitution of senses like taste and smell. If this point generalizes across many of the proper sensibles, then it seems that sensory substitution faces a serious limitation.

Even if we assume that it is possible to convey very rich information to subjects by sensory substitution devices, we might still ask whether sensory substitution has other limitations. Malika Auvray and Ophelia Deroy mentioned that sensory substitution research has not yet been able to generate a typical profile of emotional and hedonic responses. Additionally, as Jerome Dokic noted, there are some reasons to think that non-sensory perceptual feelings of familiarity and presence do not supervene on the conveyed sensory content, suggesting that such feelings may not be reliably generated by sensory substitution. As Renier pointed out, however, the absence of hedonic aspects might be due to the very basic nature of the stimuli used (lines, shapes, simple patterns etc.). Although such perceptual feelings do not supervene on the sensory content, Dokic suggested that they might be the result of a post-perceptual process which can be transferred to sensory substitution subjects. If these considerations are correct, then some limitations might only be technical ones that can be overcome in principle.

A very important feature of sensory experience is its *phenomenology*, and there seems to be a significant limitation with respect to the generation of perceptual phenomenology through sensory substitution. Several people pointed out that even if sensory substitution devices can convey a rich array of information, the *feel* of seeing something might not be transferred to other modalities. As pointed out by Macpherson, however, such worries may be motivated by an anti-representationalist assumption according to which the content of perception leaves out phenomenology. If so, then sensory substitution's limitations with respect to generating phenomenology will depend on which theories of perceptual experience are true.