

Other ways of seeing: From behavior to neural mechanisms in the online “visual” control of action with sensory substitution

Michael J. Proulx^{a,*}, James Gwinnutt^a, Sara Dell’Erba^a, Shelly Levy-Tzedek^b,
Alexandra A. de Sousa^{a,c} and David J. Brown^a

^a*Crossmodal Cognition Lab, Department of Psychology, University of Bath, Bath, UK*

^b*Cognition, Aging and Rehabilitation Lab, Recanati School for Community Health Professions,
Department of Physical Therapy & Zlotowski Center for Neuroscience, Ben Gurion University of the Negev,
Beer-Sheva, Israel*

^c*Department of Science, Bath Spa University, Bath, UK*

Abstract. Vision is the dominant sense for perception-for-action in humans and other higher primates. Advances in sight restoration now utilize the other intact senses to provide information that is normally sensed visually through sensory substitution to replace missing visual information. Sensory substitution devices translate visual information from a sensor, such as a camera or ultrasound device, into a format that the auditory or tactile systems can detect and process, so the visually impaired can see through hearing or touch. Online control of action is essential for many daily tasks such as pointing, grasping and navigating, and adapting to a sensory substitution device successfully requires extensive learning. Here we review the research on sensory substitution for vision restoration in the context of providing the means of online control for action in the blind or blindfolded. It appears that the use of sensory substitution devices utilizes the neural visual system; this suggests the hypothesis that sensory substitution draws on the same underlying mechanisms as unimpaired visual control of action. Here we review the current state of the art for sensory substitution approaches to object recognition, localization, and navigation, and the potential these approaches have for revealing a metamodal behavioral and neural basis for the online control of action.

Keywords: Sensory substitution, blindness, prosthetics, object recognition, perception for action, active sensing

Our senses allow us to interact with the world. Vision in particular is adapted to providing information about distant, silent objects by decoding the reflected or emitted light to perceive what objects are and where they are (Marr, 1982). The distinction between recognition and localization maps on to the discovery of

two processing streams in the visual brain for “where” information in the dorsal stream and “what” information in the ventral stream (Mishkin & Ungerleider, 1982). It has been suggested that these categories should be modified: the “where” category includes perception for action in addition to mere spatial perception, and would more appropriately be termed “how” rather than just “where” (Milner & Goodale, 2008). It is worth noting that, although there is this apparent

*Corresponding author: Michael J. Proulx, Department of Psychology, University of Bath, 2 South, Bath BA2 7AY, UK. Tel.: +44 1225 385963; E-mail: m.j.proulx@bath.ac.uk

dissociation of processing into two streams in the brain, there certainly must be many interactions between the two: one must know what something is to know where it is and how to act upon it. This is apparent in any experience where one simultaneously processes what and where something is while acting upon such information, but also in an extensive line of recent work that has shown that action can impact perception, and vice-versa (Biegstraaten, de Grave, Brenner, & Smeets, 2007; Cressman & Henriques, 2011). The adaptive nature of action, such as hand motion, in response to perceptual information is termed the perceptual online control of action (Gritsenko, Yakovenko, & Kalaska, 2009). Perception for action research provides a dynamic perspective given that one must continually map and re-map the surroundings in a reference frame to accommodate for self-movement and the movement of other objects (Pasqualotto & Proulx, 2012; Pasqualotto, Spiller, Jansari, & Proulx, 2013b).

For example, past research has shown that when hand motion is directed towards a visual target, an initial set of instructions is sent to the muscles controlling both hand and eye motion, and that this information is updated from retinal and extra-retinal signals as the motion occurs to fine tune the trajectory of the hand. This allows one to account for any changes in location of the target (Goodale, Pelisson, & Prablanc, 1986; Sarlegna et al., 2004) and to correct small performance errors (Gritsenko et al., 2009). This fine tuning of motion is termed online control of action and is suggested to be automatic rather than voluntary (Gritsenko et al., 2009). Transcranial magnetic stimulation (TMS) can disrupt individuals' online corrective behavior when placed over the posterior parietal cortex, suggesting it plays an important role in the online updating of motor actions (Gomi, 2008).

The online control of action can occur in the other senses, for example, humans can use auditory feedback and even echolocation to direct movements toward a target (Boyer et al., 2013; Thaler, 2013; Thaler, Arnott, & Goodale, 2011; Thaler & Goodale, 2011; Thaler, Milne, Arnott, Kish, & Goodale, 2014). Recent work to develop prosthetics for sight restoration aims to provide "visual" online control through methods of sensory substitution for vision using the other senses. Sensory substitution devices aim to provide various features and dimensions of visual information for the visually impaired by translating visual information into tactile or auditory form

(Bach-y-Rita & Kercel, 2003; Meijer, 1992; Proulx, 2010).

Here we review the research on sensory substitution for vision restoration in the context of providing the means of online control for action in the blind or blindfolded. First we will describe sensory substitution in more detail, and then we will examine how such devices are used for object recognition, localization, and action, including pointing, grasping, and navigation behaviors. We summarize with a review the neural basis for perception through such devices and note the excellent potential to reveal new, underlying neural mechanisms for the online control of action (Levy-Tzedek et al., 2012b).

1. Substituting vision with auditory or tactile perception

Vision has special properties that are challenging to convey to the other senses in that the bandwidth of vision, and the capacity for parallel processing, exceeds that of the other senses (Pasqualotto & Proulx, 2012). Parallel processing is crucial for multisensory integration and thus multisensory perception and learning. This is disrupted in congenital visual impairment because vision provides the best spatial acuity and processing capacity to integrate information from the other senses (Proulx, Brown, Pasqualotto, & Meijer, 2014). The information capacity of the human eye has been estimated to be around 4.3×10^6 bits per second (Jacobson, 1951). This is four orders of magnitude greater bandwidth than estimates of the human fingertip of 100 bits per second (Kokjer, 1987), and other areas of skin estimated even lower, from 5 to 56 bits per second (Schmidt, 1981). The information capacity of the human ear is between these two estimates, and the highest after vision, with a capacity at 10^4 bits per second (Jacobson, 1950). For this reason, although the original tactile-visual sensory substitution (TVSS) systems used the skin of the back or stomach as an analogue of the retina (Bach-y-Rita, Collins, Saunders, White, & Scadden, 1969), the modern version uses more sensitive areas such as the forehead or tongue (Bach-y-Rita & Kercel, 2003). Even so, compared to central vision it is low-resolution.

The first general-purpose auditory sensory substitution system was invented by Meijer (1992) to provide a higher resolution substitution mechanism. Although the skin/retina analogue is obvious in some ways due

to the topographic representation of spatial locations on each, the auditory system, compared to the skin, provides a higher informational capacity to convey visual information to the brain. Meijer's invention, "The vOICe" (the middle letters signify, "Oh, I see!") converts images captured by a camera into sound transmitted through headphones. Each image is translated with a left-to-right scan every one or two seconds with timing and stereo cues representing the horizontal axis, frequency representing the image's vertical axis and loudness representing pixel brightness. The input can take many forms, however it is always two-dimensional in nature, with one camera providing sufficient visual information for this. The software can run on Android phones, using the phone's camera, or with a webcam input for a PC. Spy sunglasses, with a webcam on the bridge of the nose, allow the visual input to be near the impaired or blindfolded eyes as a form of monocular vision. There are other auditory approaches to substitution. The Prosthesis for Substitution of Vision by Audition, for example, uses frequency to code for both axes by increasing from bottom to top and from left to right of the converted image, with increased representation of the center of the image as occurs with the fovea (Capelle, Trullemans, Arno, & Veraart, 1998).

Both of these devices, with extensive training, result in experiences that seem qualitatively similar to vision in some ways. Participants report distal attribution with tactile and auditory substitution for vision, rather than just merely attributing perceptual experience to the stimulated modalities (Auvray et al., 2005). Fascinatingly, a long-term user of The vOICe who acquired blindness later in life reported having visual imagery evoked by the device that is reminiscent of vision, including depth perception and smooth movement (Ward & Meijer, 2010). These findings of distal attribution and visual phenomenology converge with theoretical work concerning whether the experience of sensory substitution is more like the sensory modalities which have been substituted or those which are doing the substituting (Hurley & Noë, 2003), with some concluding that the experience is an extension of the sensory capacities of the individual to a novel sensory modality domain (Auvray & Myin, 2009). There is a unifying role for sensorimotor action for learning to use sensory substitution, and for describing the experience of it (O'Regan & Noe, 2001). Although visual information might be transformed into sound or touch, interacting with the world in an active, visual-like way

will confer something other than just the substituting sensory modality, and possibly even something resembling the substituted modality.

2. Pointing and grasping

Vision is key for being able to locate silent, distal objects and also the optimal sense to allow sensorimotor coordination for directing another person via pointing, or reaching out and grasping accurately objects within arm's reach. A few studies on blind or blindfolded subjects have examined object localization, with two conducted to also assess identification (Auvray, Hanneton, & O'Regan, 2007; Brown, Macpherson, & Ward, 2011), and one that focused specifically on perception for action (Proulx, Stoerig, Ludwig, & Knoll, 2008). Over a three week period, participants assigned to use The vOICe practiced with it daily using objects in their own homes. Proulx et al. (2008), in a study on blindfolded subjects, assessed localization with a manual search task that required pointing and responding by pressing LED lit buttons. They used a perimetry device constructed with an LED array. An LED would light up and a tone would be emitted by the controlling computer until the light had been pressed. The participants were able to generalize the self-initiated practice at home to the lab test with significant improvements in speed and accuracy. A second experiment examined the localization and grasping of natural objects placed on a large table. Although again, those trained to use The vOICe had significant improvement in locating the objects, the more impressive finding was that the participants had also learned to reach with grasp-appropriate hand configurations (Dijkerman, Milner, & Carey, 1999). Given the challenge of interpreting the sound, created by a two-dimensional, greyscale image, and determining object borders, distance, and size, this finding was surprising. Such performance implied that participants learned to extract information about size, distance, shape, and orientation, and were able to use that information to adjust grasping movements appropriately.

Research investigating the relevance of device input location has revealed how action for perception might work in non-visual ways. In the work by Proulx et al. (2008) the camera that served as the visual input capturing device for The vOICe was located near the eyes in the form of a small camera in spy sunglasses. However the camera need not be located near the eyes, nor

kept relatively static in this way. Brown, Macpherson, and Ward (2011) found that different tasks benefitted from different camera locations. Although Proulx et al. (2008) used a head-mounted camera to mimic eyesight, Auvray et al. (2007) used a handheld webcam for their experiment. Brown et al. (2011) compared experiments that required either object localization or identification with each camera position. They found an interaction of task by camera location: localization performance was superior with the glasses-mounted camera, identification performance was superior with the handheld camera. This finding implies that a sensorimotor account of learning and perception (O'Regan & Noe, 2001) applies to sensory substitution devices for the processing of "where/how", that is perception-for-action, but not necessarily for perception for object recognition ("what"). Even if using a different sense modality such as hearing, one can mimic standard perceptual-motor contingencies that are used in normal localization as revealed by having the image input near the eyes. It is important to note that the participants in the Brown et al. (2011) study were blindfolded sighted individuals. It would be interesting to see whether this result extends to not only the late, adventitiously blind who have had visual perceptual-motor experience, but also to the congenitally blind who have not had such experience. In this way the role of visual experience in development, versus innate perceptual-motor mechanisms, could be assessed similar to other work on the role of visual experience for cognition (Pasqualotto, Lam, & Proulx, 2013a; Pasqualotto & Proulx, 2012, 2013; Pasqualotto et al., 2013b; Proulx, Brown, Pasqualotto, & Meijer, 2014; Proulx et al., 2014).

These initial studies examining reaching and grasping movements guided by The vOICE sensory substitution device were ecologically valid in many ways, but did not have the ability to measure fine-grained differences in timing and accuracy. A device that has the same basic structure as The vOICE, with the addition of synthesized musical instruments as a feature to represent color, is called the EyeMusic (Levy-Tzedek, Hanassy, Abboud, Maidenbaum, & Amedi, 2012a; Levy-Tzedek, Riemer, & Amedi, 2014). These researchers examined the ability of blindfolded sighted individuals to perform fast and accurate reaching movements to targets presented by the EyeMusic, and compared them to visually guided movements. Participants were required to reach rapidly to a target. Surprisingly, participants were able to per-

form the task nearly as accurately with the device as with seeing after just a short period of training. One limitation, however, is that the participants were given a potentially unlimited amount of time to listen to the object location before the fast reaching movement was required. This delay before initiating the movement illustrates the effortful, conscious nature of using sensory substitution without extensive training, particularly in comparison to the ability of the visual system to process information rapidly and in parallel.

The crossmodal transfer of information from sensory substitution to vision (more as a version of sensory augmentation) was examined in another study by Levy-Tzedek et al. (2012b) that also used The vOICE modified to present colour with the EyeMusic. This study used a version of an online control of action visuomotor rotation task (Krakauer, Pine, Ghilardi, & Ghez, 2000; Novick & Vaadia, 2011). Participants used a joystick to control a cursor on a computer screen. Their goal was to move the cursor from a starting position to a target presented on the screen. The location of the target was alternately portrayed using vision or SSD. After several trials, the relationship between their hand movement and the location of the cursor was altered, such that the cursor movement was diverted by 30 degrees compared to the actual hand movement. In this condition, in order to have the cursor reach the target, they had to perform a skewed movement, 30 degrees away from the actual target location (for example, in order to reach a target that is 30 degrees away from the vertical, they would have to move their hand along the vertical). Importantly, the feedback suggesting that the motor command needed to be altered was provided only via the visual modality with no rotation cues given via sensory substitution. The results demonstrated that participants were quick to alter their hand movements appropriately, and that adaptation to this rotation transferred rapidly from the visual trials to the SSD trials. Aftereffects (that is, movement of the hand and the cursor in the opposite direction once rotation was removed) were found independent of the sensory modality (vision or sensory substitution via audition) in which targets were presented. This implies that the underlying neural representation of spatial information used for online action might be independent of the sensory modality with which the target is acquired. Interestingly, a transfer between vision and sensory substitution occurred with very little training, and with no conscious awareness of the presence of rotation, as reported by the participants.

The PSVA (Capelle, Trullemans, Arno, & Veraart, 1998) has also been used in localization studies. Recall that while The vOICE scans an image from left to right to create the sound, the PSVA provides a simultaneous coding of the entire image using frequency and as a result requires intentional movement to make sense of the image; this is similar to “active sensing” such as using eye movements to perceive an image. The studies described thus far with The vOICE required participants to compute depth to accurately locate and grasp the objects. A study with the PSVA investigated the problem of depth perception from monocular cues. Renier et al. (2005a) had participants locate items in depth using natural corridor cues in a simple, simulated setting. The participants were able to report accurately the depth relations between the objects in the display, though it would be of great interest to extend this work to employ reaching and grasping as well. The authors reported that all early blind participants knew that relative size was a depth cue having heard how sighted people described daily observations. This knowledge is crucial for the task because the acquisition of object identity through haptics is size invariant because any object that can be touched is perceived as its three-dimensional size. This knowledge has also been reported in studies of visual-to-tactile sensory substitution (Segond, Weiss, Kawalec, & Sampaio, 2013). Early or congenitally blind participants still have more difficulty using this knowledge than those who had visual experience previously. Learning to “see” through the other senses might require not only active experience (Proulx, 2010), but perhaps visual experience as well for full functioning or at least for the efficient acquisition of the new perceptual motor contingencies (Pasqualotto & Proulx, 2012). In fact, the initial studies with the TVSS system found that participants needed to manipulate the camera themselves in order to experience distal attribution of the scene, and that the same did not occur if the camera was manipulated by the experimenter. Therefore they needed to experience perceptual-motor contingencies via active sensing through an interaction with the object, and this form of tactile sensation that represented the image (Bach-y-Rita, Collins, Saunders, White, & Scadden, 1969).

3. Spatial navigation

Among the other senses, the auditory system has superior temporal processing (Brown & Proulx, 2013),

and as a result most studies have focused on visual-to-auditory sensory substitution devices for representing object information with high acuity and fidelity in both the spatial and temporal domains (Haigh, Brown, Meijer, & Proulx, 2013). One study also examined the use of The vOICE for spatial navigation despite the fact that image acquisition with this device occurs infrequently. Navigation clearly requires online perception and correction to avoid obstructions, yet The vOICE samples the environment with an image every 1-2 s. Such snapshots of the visual world might not be enough to allow for truly online corrections to movements, and this would be exacerbated with moving obstructions. Brown et al. (2011) had participants using The vOICE walk a short route that contained four obstacles. Participants required over five minutes to complete the task; they demonstrated improvement of over one minute within eight trials of practice.

Does navigation require the higher resolution afforded by auditory devices rather than tactile ones? Visual navigation certainly utilizes lower-resolution peripheral vision and thus perhaps a simpler representation better suits an online task such as full body navigation through space. This is a domain where visual-to-tactile sensory substitution would excel. Devices like the Tongue Display Unit (TDU), a matrix of 1-mm electrodes placed on the tongue, and conveying image information via tactile stimulation, provide constant stimulation and updating of the image input, and, like the PVSA, require movement for accurate localization and identification of objects (Matteau, Kupers, Ricciardi, Pietrini, & Pfitz, 2010). The resolution of tactile devices is lower than that of an auditory device like The vOICE. Yet, as noted, normal obstacle avoidance utilizes peripheral vision, with lower resolution compared to the fovea due to a combination of having fewer cone cells and the cortical magnification factor that favors foveal vision (Loschky, McConkie, Yang, & Miller, 2005). Peripheral vision is also a primary contributor for magnocellular processing which, like the representation provided by the TDU, is selective for contrast (the TDU has only a black and white, not greyscale, representation) and motion (Livingstone & Hubel, 1988). Thus having a sensory substitution device that provides a better representation of contrast and motion might better integrate with the neural mechanisms used for spatial navigation. Moreover assessing the importance of contrast and motion for navigation with sensory substitution would reveal whether these are the crucial computa-

tional aspects independent of sensory modality, and thus metamodal.

Second, Weiss, and Sampaio (2005) assessed spatial navigation by having the participants use the TDU while sitting, but operating a camera-carrying robot in a three-dimensional maze with a remote control. Although the accuracy and speed of the maze completion were assessed and demonstrated success from the very first attempt, it is difficult to generalize from this to real-world navigation. A more recent study with the TDU (Chebat, Schneider, Kupers, & Ptito, 2011) used a large scale navigation task in a corridor with obstacles, similar to the navigation employed with The vOICe (Brown et al., 2011). Participants were able to successfully navigate the course accurately and congenitally blind individuals were able to out-perform sighted participants, supporting other work that found route or egocentric spatial knowledge is preferred in the absence of visual experience (Pasqualotto & Proulx, 2012; Pasqualotto et al., 2013b). The participants were successful in terms of accuracy, but this study did not provide as much information about the time required to complete the task.

Navigation is a major challenge for those who have acquired blindness later in life (Sampaio & Dufier, 1988), primarily due to feelings of insecurity in relation to the environment, such as obstacles. Certainly one can learn to depend on different kinds of cues to form representations of space for successful navigation (Fortin et al., 2008) and utilize white canes and guide dogs. A recent study assessed two echolocation sensory substitution for vision devices that provide online spatial information through sound, using ultrasound emitters to detect distance from objects (Kolarik, Timmis, Cirstea, & Pardhan, 2014). They required blindfolded participants to walk through an aperture similar to a doorway, but with varied widths unknown to the user. With vision, movements were rapid and accurate, with no collisions. Two pulse-echo sensor devices were used, similar in function to sonar. The Miniguide translates obstacle distance to hand vibrations. The K-Sonar, due to having a narrower beam, translates the distance of an individual object, the existence of multiple objects, and also provides information that can support identification of objects. There were no significant differences between the two devices used. Although performance was less accurate, with some collisions, and slower, as measured by walking velocity, the study provided a proof of concept that such devices could be of use with fur-

ther training, consistent with past work with similar devices (Hughes, 2001). For example with an aperture 18% greater in width than the participant, participants had one collision every trial with K-Sonar, or every two trials with Miniguide; movement times were less than 2 seconds with vision but 5 to 8 seconds with Miniguide and K-Sonar respectively. With vision, participants rotated their shoulders far less to move through the aperture than those using the devices did. Furthermore the results implied that these echolocation devices rely on multimodal areas for depth processing in the occipito-parietal and occipito-temporal areas (Renier et al., 2005a). As demonstrated by Kolarik and colleagues, sensory substitution devices can provide another means of obstacle detection, and although some studies have examined virtual navigation (Maidenbaum, Levy-Tzedek, Chebat, & Amedi, 2013), further research examining the control of actual action based on information provided via such devices will be crucial for this form of visual rehabilitation.

4. Active sensing

Perception is necessary for action (Goodale, 2008); likewise, action plays a pivotal role in perception, in what has become known as active sensing. Active sensing has been shown to play a vital role in perception of various modalities, including the visual, auditory, tactile, and olfactory systems (Schroeder, Wilson, Radman, Scharfman, & Lakatos, 2010; Wachowiak, 2011), in humans as well as a variety of other species (Nelson & MacIver, 2006). Active sensing is the movement of the sensors for optimization of sensing performance (Mitchinson & Prescott, 2013), and it plays a central role in the process of perception (Ahissar & Arieli, 2012).

The studies reviewed here report relatively high success rates for the generation of action under SSD guidance. These results indicate that the SSD users are able to maintain the appropriate calibration of their movements based on SSD feedback, and suggests they are making use of the SSD feedback to fine-tune their movements. Movements of sighted individuals in the absence of visual feedback, have been shown to be miscalibrated, and not corrected based on proprioception alone (Levy-Tzedek, Ben Tov, & Karniel, 2011; Levy-Tzedek, Krebs, Arle, Shils, & Poizner, 2011; Soechting & Flanders, 1989; Wolpert, Ghahramani, & Jordan, 1995). It can therefore be concluded

that the effective generation of movement is a direct result of successfully deciphering cues from the SSD, rather than relying on proprioceptive cues, or on memory (e.g., from visually experiencing the experimental setup, where applicable).

Perception of depth cues plays an important role in generation of a fuller representation of spatial layouts, an ability which has been demonstrated when 2D images were conveyed via an SSD, such as the PSVA (Renier et al., 2005a). We estimate it would therefore be beneficial to combine input from SSDs that provide image information (e.g., the vOICE) with distance-estimation devices (e.g., the K-Sonar), while actively exploring a scene. Exploratory attempts in this direction have been made using newly available depth sensors such as the Kinect (Gomez, Mohammed, Bologna, & Pun, 2011).

5. Sensory substitution as cross-modal “translators”

Training on how to use SSDs is a major issue that should be considered, as the choice of a particular training paradigm may affect the ultimate performance. For example, active training has emerged as having particular benefits in recent accounts (Proulx et al., 2008; Ward & Meijer, 2010). This holds true in visual rehabilitation across the board, and applies to training on how to use retinal implants (Dagnelie, 2012) and sight-restoration surgery (Ostrovsky, Andalman, & Sinha, 2006; Putzar, Gondan, & Röder, 2012; Xu, Yu, Rowland, Stanford, & Stein, 2012). Learning to decipher visual information is a long process, requiring the learning of principles that are unique to vision, such as 3-D perspective, depth cues, variation of size with distance, transparency, reflection, color and shading effects. These are especially challenging for congenitally blind individuals, who do not have experience with this type of input. An added layer of complexity is the fact that the visual input arriving from a prosthesis or an SSD is often degraded, and harder to make sense of than a high-resolution input (Brown, Simpson, & Proulx, 2014; Haigh et al., 2013). Therefore, a carefully planned training program is of utmost importance. Combining concurrent sensory input arriving from different modalities has been shown to improve perception (for a review, see (Proulx et al., 2014)). Therefore, once the users learn how to interpret the signals arriving from the SSD, and make ‘visual’ sense of

them, they can then use the SSD in combination with other rehabilitation approaches, e.g., with a visual prosthesis, and serve as an interpreter of sorts, which can guide and assist the rehabilitation process of the more invasive approaches.

6. Using sensory substitution to reveal the neural mechanisms underlying perception for action: Ventral & dorsal streams in the blind

The study of sensory substitution and visual impairment can also reveal the neural mechanisms underlying perception for action that supersedes just one sense (Proulx et al., 2014; Ricciardi, Bonino, Pellegrini, & Pietrini, 2014); this is particularly important given that most research in this area falls within the domain of vision. Visual information is segregated into different features early in processing, even at the level of the retina, and later into two parallel cortical pathways, the ventral stream and the dorsal stream, which appear to be functionally segregated, with the former discriminating object features and the later handling information about spatial location (Goodale, 2008). The ventral occipitotemporal “what” pathway extends from the primary visual cortex (V1) through visual area 4 (V4) to inferotemporal areas and then up to the prefrontal cortex. The dorsal occipitoparietal “where/how” pathway from V1 leads through the posterior parietal cortex towards the premotor cortex (Striem-Amit, Dakwar, Reich, & Amedi, 2012). The dorsal stream is critical for real time control of action, providing visual guidance for movement (Milner, 2012). The ventral stream supports object, event, and action identification (Goodale, 2008) as well as feature analysis for object identity independent of variance in other features for which identity should be invariant, such as position and orientation (Freeman & Ziemba, 2011). The role of the ventral stream in action, then, is to provide visual information to enable the identification of a specific object and activate the cognitive systems to plan the action performed towards the object. These streams were primarily seen as visual because deficits caused by lesions to these areas cause visual aberrations (Milner & Goodale, 2008).

A growing body of research examining perception, action, and cognition in the visually impaired suggests that the brain is not necessarily segregated into sensory-specific areas but instead is metamodal and organized

in terms of the tasks, functions, and computations that the areas carry out (Pascual-Leone & Hamilton, 2001; Proulx et al., 2014); this is sometimes referred to as supramodal as a synonym for metamodal in the literature (Kupers, Pietrini, Ricciardi, & Ptito, 2011; Kupers & Ptito, 2011; Ricciardi & Pietrini, 2011). This perspective, in contrast with the classical one, classifies brain areas by the form of information processing carried out (e.g., shape recognition) independent of the sensory modality that appears to be employed. First we will review the evidence for the metamodal organization of the brain. Then we will note the natural extension of this form of functional organization to perception for action, in much the same way that the dorsal stream has been re-characterized from processing “where” information (Mishkin & Ungerleider, 1982), to being more task-specified as the processing of “how” (Milner & Goodale, 2008) such information can be acted upon.

Two examples of cortical regions, first thought to be visual in nature, but later discovered to be metamodal are the fusiform face area (Kanwisher, McDermott, & Chun, 1997) and the parahippocampal place area (Epstein & Kanwisher, 1998) that are now known to respond to the same categories even when only perceived by touch or sound (Ricciardi & Pietrini, 2011). An extensive review by Ricciardi and Pietrini (2011) details many sources of evidence for the supramodal, or metamodal, functional architecture of the brain.

Striem-Amit et al. (2012) conducted a brain imaging study to examine the functional segregation of the dorsal and ventral streams in the congenitally blind. Using The vOICE, congenitally blind participants were presented different simple novel shapes in different areas of the “visual” field, and these were transformed into sound. The participants were asked either to identify or locate the shape. This elicited clear differentiation between ventral and dorsal streams, indicating that even with no visual experience, these two streams still encode the tasks they would be expected to perform had it been visual information. To be more specific, shape identity tasks activated the ventral occipital temporal sulcus, which is located in the midst of the ventral stream. In contrast, the localization task preferentially activated auditory regions, such as the supramarginal gyrus located in the inferior parietal lobe, as well as the precuneus, a higher order area of the dorsal stream.

Criticisms that are levelled at studies using SSDs are that they usually study highly proficient users of the system, who may have already undergone a plastic

change, such as Ptito, Moesgaard, Gjedde, and Kupers (2005) who showed increased occipital cortex activation following SSD training. Striem-Amit et al.’s (2012) study circumvents this limitation by only training participants to use the devices for a maximum of one and a half hours and participants were completely naïve to The vOICE before this training, thus avoiding the problem that brain imaging might be measuring an artificial plastic change rather than what is inherent in the congenitally blind.

Furthermore, many studies exploring the use of SSDs in the congenitally blind use familiar, well-practiced stimuli, such as Kim and Zatorre (2011) who trained their participants for five days with the stimuli used for the experiment. This means that potential findings were confounded by participants using memory of the stimuli presented in training, rather than shape processing in the occipital region. Striem-Amit et al. (2012) used novel stimuli to ensure that memory was not playing a role in the activation of dorsal and ventral streams, just shape and location processing. The neural correlates of action control and representation has not been only studied using sensory substitution, but there is extensive evidence using standard paradigms in sensory-deprived individuals, further giving support to the metamodal nature of the underlying neural architecture for action (Renzi et al., 2013; Emiliano Ricciardi et al., 2009; Röder, Kusmirek, Spence, & Schicke, 2007).

Qin and Yu (2013) postulated that the metamodal nature of the visual cortex could be due to a plastic change in the brain that is caused by the deprivation of visual stimuli. Thus, rather than being inherently metamodal, the visual cortex has been shown to be recruited by other higher functioning processes, such as memory and cognition in the blind (Burton, Sinclair, & Agato, 2012; Renier et al., 2010), and therefore might not be metamodal naturally. Striem-Amit et al. (2012) refute this suggestion by demonstrating this metamodality in the blindfolded sighted as well as the blind, indicating that the metamodal nature of the two streams is not due to sensory deprivation.

This study seems to suggest that a metamodal approach to understanding the brain structure is a more valid approach compared with the standard unimodal model. If the metamodal theory is to be substantiated, more areas of the brain must be demonstrated to be metamodal. Reich, Szwed, Cohen, and Amedi (2011) studied the “visual” word form area (VWFA), a component of the visual cortex, to try to establish whether this

area is metamodal. The VWFA in the sighted is activated across writing systems and scripts and encodes letter strings irrespective of case, font or location in the visual field. In this study blind participants were presented with Braille words and nonsense Braille letter strings. The authors used fMRI to investigate whether the VWFA would be activated in response to the tactile letters. The investigators found a striking differentiation between the real Braille words and the nonsense words in the congenitally blind, indicating that the VWFA is activated for reading words regardless of input, plus develops specialization for reading regardless of visual experience, supporting the metamodal theory. Braille reading activating the primary visual cortex in the blind has also been shown by Cohen et al. (1997), thus strengthening the idea that the visual cortex is a metamodal area of the brain that can be activated by any sensory modality.

A further brain area that has been put forward as metamodal is hMT+ (also called the fifth visual area, V5), which encodes for motion. A study by Matteau et al. (2010) found this area was activated when perceiving motion using an SSD that rests on the tongue in both the sighted and the blind, again indicating that this area encodes information about motion regardless of sensory input.

So far in this review only visual areas have been discussed as being metamodal. For the theory to be substantiated fully, other areas of the brain that have traditionally been considered to be unimodal must be demonstrated to be activated by more than one modality. For example, the auditory cortex would be expected to be activated if a non-audio input was presented that required the same computation as an audio signal. Calvert et al. (1997) used an fMRI with participants with normal hearing. The authors first identified the auditory cortex by playing participants speech through headphones and then asked them to repeat the words back to themselves in their heads. Then they played video of a face silently mouthing numbers, which the participants had to repeat back to themselves in their heads. This lip-reading led to activation in the auditory cortex, despite there being no auditory signal present, indicating that the auditory cortex is also metamodal (Calvert & Thesen, 2004).

This is further corroborated by Okada et al. (2010) who presented 20 participants with audio input alone compared with audio and an accompanying congruent visual speech input. Audio plus visual led to more activation in the auditory cortex than audio input alone,

indicating that areas of the so-called auditory cortex are similarly responsive to visual stimuli, again supporting the idea of the brain as a series of modules that deal with specific computations, rather than being described by the input they receive. Other research suggests similar findings for other “primary” sensory areas, such as the primary olfactory cortex, which has been shown to be responsive to other associative information (Chapuis & Wilson, 2012; Weiss & Sobel, 2012). Note that the metamodal, or supramodal, nature of a brain area can be further demonstrated by studies comparing sighted and early blind individuals, as this would exclude visual imagery as a potential confound. Indeed many of these areas described have been tested with such an experimental design, and a recent review (Ricciardi et al., 2014) provides examples of such evidence.

How might these cortical regions, thought to be primary sensory areas, respond to information provided by other sensory modalities? For example, the visual cortex must be connected to the cochlear nuclei in the brain stem (Recanzone & Sutter, 2008) for there to be activation in this area when presented with images in sound via The vOICe. Two hypotheses have been put forward in explanation of how these, previously thought to be unlinked, peripheral sensors and brain regions are connected. According to the *cortical reorganization* hypothesis, cross modal brain responses are mediated by the formation of new pathways in the sensory deprived brain (Bronchti et al., 2002). A more compelling argument is the *unmasking* hypothesis, which states that loss of sensory input induces unmasking and strengthening of already existing neuronal connections (Kupers & Ptito, 2011).

The unmasking hypothesis is suggested here to be correct, as blind participants in the Striem-Amit et al. (2012) study showed activation in visual pathways when using an SSD with only minimal training, precluding the idea that the brain had a chance to rewire itself, instead seeing the connections as extant and unmasked by these types of experiment. This is further substantiated by Striem-Amit et al. (2012) who also showed the same activation in sighted blind-folded participants as in the blind, indicating that it is not a plastic change that occurred due to deprivation of vision.

Furthermore Falchier, Clavagnier, Barone, and Kennedy (2002), using retrograde tracers on cynomolgus monkeys, showed that the primary visual cortex (V1) receives input not only from the visual thalamus but also receives projections from the auditory cortex as well as polysensory areas of the temporal

lobe. Further anatomical evidence of multiple other connections projecting to the visual cortex comes from Rockland and Ojima (2003), who showed that areas V1 and secondary visual cortex (V2) receive projections from several parietal and auditory association areas in macaque monkeys. These anatomical explorations in monkeys show that the previously thought of as unimodal areas of the brain are actually connected to other sensory areas directly, rather than just via upstream multimodal association areas, and these data support the idea that the connections are extant in all and are unmasked due to sensory substitution tasks rather than created by the brain following sensory deprivation.

However the debate continues, with evidence coming from Kupers et al. (2006) who used TMS on blind and sighted participants trained to use the tongue display unit (TDU). The TDU is a sensory substitution device that places a 3×3 cm electrode array consisting of 144 gold plates on the tongue, and an electrical pulse is generated by each electrode to correspond to a pixel on the image that is being assessed (Bach-y-Rita & Kercel, 2003; Ptito et al., 2005). When TMS over the occipital cortex was used, blind participants reported somatotopically organized tactile sensations on the tongue, whereas sighted individuals who had also been trained with the TDU did not report this (Kupers et al., 2006). This indicates that a few alternative explanations might be possible. First, that there are differences in the blind that are not found in the sighted, suggesting that instead of there being masked connections between peripheral tactile sensors and the occipital cortex in everyone, there is actually a new connection created between cortex and sensor in the sensory deprived, supporting the cortical reorganization hypothesis. Another option is that the unmasking occurs to a different extent for blind and sighted individuals. That is, perhaps in the blind brain there is the potential for faster, more efficient unmasking than in the sighted brain.

Based on the evidence presented here, areas of the visual cortex have been demonstrated to be metamodal in nature, such as the dorsal and ventral streams, the VWFA and V5. Further the auditory cortex has been shown to become active when an appropriate computational task is presented to it, despite this task coming in the form of visual input (Calvert et al., 1997). Therefore a system which describes brain structure based on the task or computations that an area performs, rather than the sensory modality that is most often encoded by an area, seems to be a more accurate taxonomy. Previ-

ously the brain areas were thought to only be connected to the specific sensory modalities that they computed. However, evidence from lesion studies in monkeys and brain imaging studies in humans have shown that these metamodal areas are in fact linked up to peripheral sensors that were previously not thought to project to them. This leads to the idea that the brain is split into metamodal, computational based modules, that are linked to all sensory inputs and that can process information that is requisite to their computational task, regardless of the modality that this information originates from.

The results derived from blind participants supporting a metamodal organization of the brain extend to sighted participants as well. Long-term blindfolding of sighted participants with training in using the vOICE (Proulx et al., 2008) resulted in brain activation similar to that of blind participants in visual cortex but with auditory stimulation (Boroojerdi et al., 2000; Facchini & Aglioti, 2003; Lewald, 2007; Merabet et al., 2008; Pollok, Schnitzler, Stoerig, Mierdorf, & Schnitzler, 2005). Such effects can occur in a matter of days, too, as in a study with five days of visual deprivation and Braille training (Kauffman, Theoret, & Pascual-Leone, 2002; Pascual-Leone & Hamilton, 2001). There is mounting evidence that the visual cortex is functionally involved in auditory and tactile tasks in sighted subjects as well (Driver & Noesselt, 2008; Ghazanfar & Schroeder, 2006; Liang, Mouraux, Hu, & Iannetti, 2013; Sathian & Zangaladze, 2002; Zangaladze, Epstein, Grafton, & Sathian, 1999).

Thus far, research using sensory substitution has revealed the metamodal properties of brain areas once thought to be visual in nature. Although such research has targeted both ventral and dorsal streams of processing, those studies examining the dorsal stream have focused on “where” information in the perceptual sense rather than “how” information in the perception for action sense (Striem-Amit et al., 2012). Much of what is known about the neural basis of perception for action comes from studies of vision. For example, a large body of work has detailed the role of the anterior intra-parietal area (AIP) in macaques for the processing of object shape, size and orientation to guide grasping (Taira, Mine, Georgopoulos, Murata, & Sakata, 1990). A similar region has been found in humans as well (Binkofski et al., 1998), with a number of studies detailing the role of the posterior parietal cortex for the coordination of eye and hand movements when reaching for targets (Desmurget et al., 1999; Van Donkelaar, Lee, & Drew, 2000). Are these vision-specific percep-

tion for action areas, or might non-visual information guide hand movements and grasping in these cortical regions? The metamodal hypothesis would have one predict that the task (grasping and hand movements) of AIP defines its function, and that any sensory input that provides the necessary spatial information and object features would also evoke such functional recruitment. Sensory substitution would be an ideal approach to examine whether AIP and other areas involved in perception for action are not specific to one sensory modality, vision, but rather allow for precise hand and grasping movements on the basis of any informative sensory modality that provides input.

7. Shared spatial representation across the senses

Spatial information is shared across sensory modalities, such that information obtained via one modality (e.g., vision) is utilized when making movements based on input from another sensory modality (e.g., audition) (Levy-Tzedek et al., 2012b). These findings are in line with evidence from neural recordings showing that the dorsal stream participates in mediation of sensory-motor control, even when the sensory modality is not vision (Fiehler, Burke, Engel, Bien, & Rösler, 2008), suggesting a modality-independent representation of action control.

Several studies have demonstrated (Rossetti, Stelmach, Desmurget, Prablanc, & Jeannerod, 1994) that direct vision of the target immediately prior to movement significantly increased movement accuracy, whereas a delay between target presentation and movement resulted in deteriorated accuracy; that is to say, the longer participants had to use their cognitive representation of the target, the less accurate were their movements; this line of results was the basis for the suggestion that vision is needed to calibrate the proprioceptive map; and yet, participants in the studies reported here were able to use the auditory or tactile information (possibly to create a ‘visual’-like representation which is the ultimate goal of SSDs for the blind), in some cases without the possibility to perform vision-based calibration between trials, to both plan the movements and improve the accuracy on subsequent trials.

Proulx and colleagues (2014) reviewed these and other results to suggest that the representation of space might not be dependent on the modality through which

the spatial information is received. This is supported by brain-imaging studies indicating that areas within the visual dorsal stream, such as the precuneus (Renier & De Volder, 2010) and the right anterior middle occipital gyrus (Collignon, Champoux, Voss, & Lepore, 2011), showed a functional preference for spatial information delivered using the auditory or tactile modalities in congenitally blind and in sighted individuals. These brain areas are also activated when sound encodes spatial information, rather than arriving from different spatial locations, in both sighted and congenitally blind individuals (Striem-Amit et al., 2012). This hypothesis, of a task-specific, rather than a modality-specific structure of brain areas was similarly suggested by research indicating that, during a spatial navigation task with the TDU, congenitally blind individuals were found to recruit the posterior parahippocampus and posterior parietal and ventromedial occipito-temporal cortices, areas that are involved in spatial navigation under full vision (Chebat et al., 2011); similarly, the dorsal pathway has been shown to mediate movement control when only kinesthetic, and no visual information was used to guide actions even when tested in congenitally blind individuals, with no prior visual experience (Fiehler et al., 2008).

It has been suggested that blind individuals tend to rely on a more egocentric, rather than an allocentric frame of reference (Cattaneo, Bhatt, Merabet, Pece, & Vecchi, 2008; Pasqualotto & Proulx, 2012; Pasqualotto et al., 2013b; Röder et al., 2007). As data are accumulated, supporting the hypothesis that the representation of space in the brain is modality-independent, or that very little training is required to create a ‘visual’-like representation of space using sounds, we suggest that SSDs, by providing detailed spatial information for people with impaired vision, might allow them to employ an allocentric spatial frame of reference when interacting with their environment.

8. Future directions for the online visual control of action through sensory substitution

Here we have reviewed the few studies that have begun to explore how sensory substitution might allow for the online control of action via visual information perceived through sound or touch. One limitation of many of these studies so far is that performance has often been assessed solely in terms of accuracy, with

only a few that have also reported response times so that the full efficiency of the behaviors could be assessed.

There are a number of areas where the classic approaches to studying the online control of action can be adopted to address interesting issues with sensory substitution. One fascinating example would be to examine how size illusions via sensory substitution influence perception and action, and whether visual experience is necessary for an illusion to be effective. A crucial test to assess whether there is a clear distinction between perception and action has been whether a size illusion, such as the Müller-Lyer illusion, impacts both perceived size and grasping. Recent research (Biegstraaten et al., 2007), and a comprehensive review (Bruno & Franz, 2009), suggest that there is not necessarily a dissociation between perception and action. This has also been supported by work in other, related areas using illusions and other approaches (Bruno, Knox, & de Grave, 2010; de Brouwer, Brenner, Medendorp, & Smeets, 2014; de Grave, Brenner, & Smeets, 2004; Foster, Kleinholtermann, Leifheit, & Franz, 2012; Franz, Hesse, & Kollath, 2009; Mendoza, Elliott, Meegan, Lyons, & Welsh, 2006). This would be an interesting approach to assess the mechanisms of online control with sensory substitution because the perception of size using other size illusions has also been conducted with sensory substitution. Laurent Renier and colleagues used the PSVA sensory substitution device to assess whether blind and sighted participants would perceive the illusion through interpretation of the sound created by the device (Renier et al., 2005b; Laurent Renier, Bruyer, & Volder, 2006). In these studies they found that although the sighted blindfolded and late blind participants perceived the Ponzo illusion and the Vertical-Horizontal illusions just as the sighted do, the early blind participants did not. This has the fascinating implication that visual experience is necessary to have the expectations required to perceive size in a way that is influenced by context. They did not assess perception versus action, and it would be of great interest to assess whether these illusions, perceived through sensory substitution, affect grasping differentially depending on visual experience. Another interesting extension would be to assess the impact of using another sensory substitution approach, such as tactile substitution, on the influence of size illusions. Given the fascinating finding in other work that the online control of action might be sensory modality independent as seen from converging work looking at learning and adaptation with sensory sub-

stitution devices (Levy-Tzedek et al., 2012b), there is a lot of potential to use sensory substitution as a means for revealing the underlying mechanisms for the perceptual control of motor action. The next step would be to move beyond the limited data available on online control with full reporting of both response time and accuracy to allow a full assessment of sensory substitution performance. Future work looking at having shifting targets at the time of response initiation and double-step paradigms for reaching to targets (Goodale et al., 1986; Gritsenko et al., 2009) would best allow work on perception for action with sensory substitution to contribute to the fundamental understanding of online control.

Acknowledgments

This work was supported by a grant from the EPSRC to Michael J. Proulx (EP/J017205/1). Partial support was provided for Shelly Levy-Tzedek by the Helmsley Charitable Trust through the Agricultural, Biological and Cognitive Robotics Center of Ben-Gurion University of the Negev.

References

- Ahissar, Ehud, & Arieli, Amos. (2012). Seeing via miniature eye movements: A dynamic hypothesis for vision. *Frontiers in Computational Neuroscience*, 6.
- Auvray, M., Hannequin, S., & O'Regan, J.K. (2007). Learning to perceive with a visuo-auditory substitution system: Localisation and object recognition with 'the Voice'. *Perception*, 36(3), 416-430.
- Auvray, M., & Myin, E. (2009). Perception with compensatory devices: From sensory substitution to sensorimotor extension. *Cogn Sci*, 33(6), 1036-1058. doi: 10.1111/j.1551-6709.2009.01040.x
- Bach-y-Rita, P., Collins, C.C., Saunders, F.A., White, B., & Scadden, L. (1969). Vision substitution by tactile image projection. *Nature*, 221(5184), 963-964.
- Bach-y-Rita, P., & Kercel, S.W. (2003). Sensory substitution and the human-machine interface. *Trends Cogn Sci*, 7(12), 541-546. doi: S1364661303002900 [pii]
- Biegstraaten, M., de Grave, D.D., Brenner, E., & Smeets, J.B. (2007). Grasping the Muller-Lyer illusion: Not a change in perceived length. *Exp Brain Res*, 176(3), 497-503. doi: 10.1007/s00221-006-0744-8
- Binkofski, F., Dohle, C., Posse, S., Stephan, K.M., Hefter, H., Seitz, R.J., & Freund, H.J. (1998). Human anterior intraparietal area subserves prehension: A combined lesion and functional MRI activation study. *Neurology*, 50(5), 1253-1259.

- Boroojerdi, B., Bushara, K.O., Corwell, B., Immisch, I., Battaglia, F., Muellbacher, W., & Cohen, L.G. (2000). Enhanced excitability of the human visual cortex induced by short-term light deprivation. *Cereb Cortex*, *10*(5), 529-534.
- Boyer, E.O., Babayan, B.M., Bevilacqua, F., Noisternig, M., Warusfel, O., Roby-Brami, A. (2013). From ear to hand: The role of the auditory-motor loop in pointing to an auditory source. *Front Comput Neurosci*, *7*, 26. doi: 10.3389/fncom.2013.00026
- Bronchti, G., Heil, P., Sadka, R., Hess, A., Scheich, H., & Wollberg, Z. (2002). Auditory activation of "visual" cortical areas in the blind mole rat (*Spalax ehrenbergi*). *Eur J Neurosci*, *16*(2), 311-329.
- Brown, D.J., Simpson, A.J., & Proulx, M.J. (2014). Visual objects in the auditory system in sensory substitution: How much information do we need? *Multisens Res*, *27*(5-6), 337-357.
- Brown, D., Macpherson, T., & Ward, J. (2011). Seeing with sound? exploring different characteristics of a visual-to-auditory sensory substitution device. *Perception*, *40*(9), 1120-1135.
- Brown, David J., & Proulx, Michael J. (2013). Increased Signal Complexity Improves the Breadth of Generalization in Auditory Perceptual Learning. *Neural Plasticity*, *2013*, 9. doi: 10.1155/2013/879047
- Bruno, N., & Franz, V.H. (2009). When is grasping affected by the Muller-Lyer illusion? A quantitative review. *Neuropsychologia*, *47*(6), 1421-1433. doi: S0028-3932(08)00419-3 [pii] 10.1016/j.neuropsychologia.2008.10.031
- Bruno, N., Knox, P.C., & de Grave, D.D. (2010). A meta-analysis of the effect of the Muller-Lyer illusion on saccadic eye movements: No general support for a dissociation of perception and oculomotor action. *Vision Res*, *50*(24), 2671-2682. doi: S0042-6989(10)00461-X [pii] 10.1016/j.visres.2010.09.016
- Burton, H., Sinclair, R.J., & Agato, A. (2012). Recognition memory for Braille or spoken words: An fMRI study in early blind. *Brain Res*, *1438*, 22-34. doi: 10.1016/j.brainres.2011.12.032
- Calvert, G.A., Bullmore, E.T., Brammer, M.J., Campbell, R., Williams, S.C., McGuire, P.K. (1997). Activation of auditory cortex during silent lipreading. *Science*, *276*(5312), 593-596.
- Calvert, G.A., & Thesen, T. (2004). Multisensory integration: Methodological approaches and emerging principles in the human brain. *J Physiol Paris*, *98*(1-3), 191-205. doi: 10.1016/j.jphysparis.2004.03.018
- Capelle, C., Trullemans, C., Arno, P., & Veraart, C. (1998). A real-time experimental prototype for enhancement of vision rehabilitation using auditory substitution. *IEEE Trans Biomed Eng*, *45*(10), 1279-1293. doi: 10.1109/10.720206
- Cattaneo, Z., Bhatt, E., Merabet, L.B., Pece, A., & Vecchi, T. (2008). The influence of reduced visual acuity on age-related decline in spatial working memory: An investigation. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn*, *15*(6), 687-702. doi: 10.1080/13825580802036951
- Chapuis, Julie, & Wilson, Donald A. (2012). Bidirectional plasticity of cortical pattern recognition and behavioral sensory acuity. *Nat Neurosci*, *15*(1), 155-161. doi: <http://www.nature.com/neuro/journal/v15/n1/abs/nn.2966.html#supplementary-information>
- Chebat, D.R., Schneider, F.C., Kupers, R., & Ptito, M. (2011). Navigation with a sensory substitution device in congenitally blind individuals. *Neuroreport*, *22*(7), 342-347. doi: 10.1097/WNR.0b013e3283462def
- Cohen, L.G., Celnik, P., Pascual-Leone, A., Corwell, B., Falz, L., Dambrosia, J. (1997). Functional relevance of cross-modal plasticity in blind humans. *Nature*, *389*(6647), 180-183. doi: 10.1038/38278
- Collignon, O., Champoux, F., Voss, P., & Lepore, F. (2011). Sensory rehabilitation in the plastic brain. *Prog Brain Res*, *191*, 211-231. doi: 10.1016/B978-0-444-53752-2.00003-5
- Cressman, E.K., & Henriques, D.Y. (2011). Motor adaptation and proprioceptive recalibration. *Prog Brain Res*, *191*, 91-99. doi: 10.1016/B978-0-444-53752-2.00011-4
- Dagnelie, G. (2012). Retinal implants: Emergence of a multidisciplinary field. *Curr Opin Neurol*, *25*(1), 67-75. doi: 10.1097/WCO.0b013e32834f02c3
- de Brouwer, A.J., Brenner, E., Medendorp, W.P., & Smeets, J.B. (2014). Time course of the effect of the Muller-Lyer illusion on saccades and perceptual judgments. *J Vis*, *14*(1). doi: 10.1167/14.1.4
- de Grave, D.D., Brenner, E., & Smeets, J.B. (2004). Illusions as a tool to study the coding of pointing movements. *Exp Brain Res*, *155*(1), 56-62. doi: 10.1007/s00221-003-1708-x
- Desmurget, M., Epstein, C.M., Turner, R.S., Prablanc, C., Alexander, G.E., & Grafton, S.T. (1999). Role of the posterior parietal cortex in updating reaching movements to a visual target. *Nat Neurosci*, *2*(6), 563-567. doi: 10.1038/9219
- Dijkerman, H.C., Milner, A.D., & Carey, D.P. (1999). Motion parallax enables depth processing for action in a visual form agnostic when binocular vision is unavailable. *Neuropsychologia*, *37*(13), 1505-1510.
- Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron*, *57*(1), 11-23. doi: 10.1016/j.neuron.2007.12.013
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598-601. doi: 10.1038/33402
- Facchini, S., & Aglioti, S.M. (2003). Short term light deprivation increases tactile spatial acuity in humans. *Neurology*, *60*(12), 1998-1999.
- Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *J Neurosci*, *22*(13), 5749-5759. doi: 20026562
- Fiehler, Katja, Burke, Michael, Engel, Annerose, Bien, Siegfried, & Rösler, Frank. (2008). Kinesthetic working memory and action control within the dorsal stream. *Cerebral Cortex*, *18*(2), 243-253.
- Fortin, M., Voss, P., Lord, C., Lassonde, M., Pruessner, J., Saint-Amour, D. (2008). Wayfinding in the blind: Larger hippocampal volume and supranormal spatial navigation. *Brain*, *131*(Pt 11), 2995-3005. doi: 10.1093/brain/awn250
- Foster, R.M., Kleinholdermann, U., Leifheit, S., & Franz, V.H. (2012). Does bimanual grasping of the Muller-Lyer illusion provide evidence for a functional segregation of dorsal and

- ventral streams? *Neuropsychologia*, 50(14), 3392-3402. doi: 10.1016/j.neuropsychologia.2012.09.026
- Franz, V.H., Hesse, C., & Kollath, S. (2009). Visual illusions, delayed grasping, and memory: No shift from dorsal to ventral control. *Neuropsychologia*, 47(6), 1518-1531. doi: 10.1016/j.neuropsychologia.2008.08.029
- Freeman, J., & Ziemba, C.M. (2011). Unwrapping the ventral stream. *J Neurosci*, 31(7), 2349-2351. doi: 10.1523/JNEUROSCI.6191-10.2011
- Ghazanfar, A.A., & Schroeder, C.E. (2006). Is neocortex essentially multisensory? *Trends Cogn Sci*, 10(6), 278-285. doi: DOI 10.1016/j.tics.2006.04.008
- Gomez, Juan Diego, Mohammed, Sinan, Bologna, Guido, & Pun, Thierry. (2011). *Toward 3D scene understanding via audio-description: Kinect-iPad fusion for the visually impaired*. Paper presented at the The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility.
- Gomi, H. (2008). Implicit online corrections of reaching movements. *Curr Opin Neurobiol*, 18(6), 558-564. doi: 10.1016/j.conb.2008.11.002
- Goodale, M.A. (2008). Action without perception in human vision. *Cogn Neuropsychol*, 25(7-8), 891-919. doi: 10.1080/02643290801961984
- Goodale, M.A., Pelisson, D., & Prablanc, C. (1986). Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature*, 320(6064), 748-750. doi: 10.1038/320748a0
- Gritsenko, V., Yakovenko, S., & Kalaska, J.F. (2009). Integration of predictive feedforward and sensory feedback signals for online control of visually guided movement. *J Neurophysiol*, 102(2), 914-930. doi: 10.1152/jn.91324.2008
- Haigh, A., Brown, D.J., Meijer, P., & Proulx, M.J. (2013). How well do you see what you hear? The acuity of visual-to-auditory sensory substitution. *Front Psychol*, 4, 330. doi: 10.3389/fpsyg.2013.00330
- Hughes, B. (2001). Active artificial echolocation and the nonvisual perception of aperture passability. *Hum Mov Sci*, 20(4-5), 371-400.
- Hurley, Susan, & Noë, Alva. (2003). Neural Plasticity and Consciousness. *Biology and Philosophy*, 18(1), 131-168.
- Kanwisher, N., McDermott, J., & Chun, M.M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *J Neurosci*, 17(11), 4302-4311.
- Kauffman, T., Theoret, H., & Pascual-Leone, A. (2002). Braille character discrimination in blindfolded human subjects. *Neuroreport*, 13(5), 571-574.
- Kim, J.K., & Zatorre, R.J. (2011). Tactile-auditory shape learning engages the lateral occipital complex. *J Neurosci*, 31(21), 7848-7856. doi: 10.1523/JNEUROSCI.3399-10.2011
- Kokjer, K.J. (1987). The Information Capacity of the Human Fingertip. *Ieee Transactions on Systems Man and Cybernetics*, 17(1), 100-102. doi: Doi 10.1109/Tsmc.1987.289337
- Kolarik, A.J., Timmis, M.A., Cirstea, S., & Pardhan, S. (2014). Sensory substitution information informs locomotor adjustments when walking through apertures. *Exp Brain Res*, 232(3), 975-984. doi: 10.1007/s00221-013-3809-5
- Krakauer, J.W., Pine, Z.M., Ghilardi, M.F., & Ghez, C. (2000). Learning of visuomotor transformations for vectorial planning of reaching trajectories. *J Neurosci*, 20(23), 8916-8924.
- Kupers, R., Fumal, A., de Noordhout, A.M., Gjedde, A., Schoenen, J., & Ptito, M. (2006). Transcranial magnetic stimulation of the visual cortex induces somatotopically organized qualia in blind subjects. *Proc Natl Acad Sci U S A*, 103(35), 13256-13260. doi: 10.1073/pnas.0602925103
- Kupers, R., Pietrini, P., Ricciardi, E., & Ptito, M. (2011). The nature of consciousness in the visually deprived brain. *Front Psychol*, 2, 19. doi: 10.3389/fpsyg.2011.00019
- Kupers, R., & Ptito, M. (2011). Insights from darkness: What the study of blindness has taught us about brain structure and function. *Prog Brain Res*, 192, 17-31. doi: 10.1016/B978-0-444-53355-5.00002-6
- Levy-Tzedek, S., Ben Tov, M., & Karniel, A. (2011). Rhythmic movements are larger and faster but with the same frequency on removal of visual feedback. *J Neurophysiol*, 106(5), 2120-2126. doi: 10.1152/jn.00266.2011
- Levy-Tzedek, S., Hanassy, S., Abboud, S., Maidenbaum, S., & Amedi, A. (2012a). Fast, accurate reaching movements with a visual-to-auditory sensory substitution device. *Restor Neurol Neurosci*, 30(4), 313-323. doi: 10.3233/RNN-2012-110219
- Levy-Tzedek, S., Krebs, H.I., Arle, J.E., Shils, J.L., & Poizner, H. (2011). Rhythmic movement in Parkinson's disease: Effects of visual feedback and medication state. *Exp Brain Res*, 211(2), 277-286. doi: 10.1007/s00221-011-2685-0
- Levy-Tzedek, S., Novick, I., Arbel, R., Abboud, S., Maidenbaum, S., Vaadia, E., & Amedi, A. (2012b). Cross-sensory transfer of sensory-motor information: Visuomotor learning affects performance on an audiomotor task, using sensory-substitution. *Sci Rep*, 2, 949. doi: 10.1038/srep00949
- Levy-Tzedek, S., Riemer, D., & Amedi, A. (2014). Color improves "visual" acuity via sound. *Frontiers in Neuroscience*, 8, 358. <http://doi.org/10.3389/fnins.2014.00358>
- Lewald, J. (2007). More accurate sound localization induced by short-term light deprivation. *Neuropsychologia*, 45(6), 1215-1222. doi: 10.1016/j.neuropsychologia.2006.10.006
- Liang, M., Mouraux, A., Hu, L., & Iannetti, G.D. (2013). Primary sensory cortices contain distinguishable spatial patterns of activity for each sense. *Nat Commun*, 4, 1979. doi: 10.1038/ncomms2979
- Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, 240(4853), 740-749. doi: 10.1126/science.3283936
- Loschky, Lester, McConkie, George, Yang, Jian, & Miller, Michael. (2005). The limits of visual resolution in natural scene viewing. *Vis Cogn*, 12(6), 1057-1092. doi: 10.1080/13506280444000652
- Maidenbaum, S., Levy-Tzedek, S., Chebat, D.R., & Amedi, A. (2013). Increasing Accessibility to the Blind of Virtual Environments, Using a Virtual Mobility Aid Based On the "EyeCane": Feasibility Study. *Plos One*, 8(8). doi: ARTN e72555 DOI 10.1371/journal.pone.0072555

- Marr, David. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W.H. Freeman.
- Matteau, I., Kupers, R., Ricciardi, E., Pietrini, P., & Ptito, M. (2010). Beyond visual, aural and haptic movement perception: HMT+ is activated by electrotactile motion stimulation of the tongue in sighted and in congenitally blind individuals. *Brain Res Bull*, 82(5-6), 264-270. doi: 10.1016/j.brainresbull.2010.05.001
- Meijer, P.B. (1992). An experimental system for auditory image representations. *IEEE Trans Biomed Eng*, 39(2), 112-121. doi: 10.1109/10.121642
- Mendoza, J.E., Elliott, D., Meegan, D.V., Lyons, J.L., & Welsh, T.N. (2006). The effect of the Muller-Lyer illusion on the planning and control of manual aiming movements. *J Exp Psychol Hum Percept Perform*, 32(2), 413-422. doi: 2006-04604-015 [pii] 10.1037/0096-1523.32.2.413
- Merabet, L.B., Hamilton, R., Schlaug, G., Swisher, J.D., Kiriakopoulos, E.T., Pitskel, N.B. (2008). Rapid and reversible recruitment of early visual cortex for touch. *PLoS One*, 3(8), e3046. doi: 10.1371/journal.pone.0003046
- Milner, A.D. (2012). Is visual processing in the dorsal stream accessible to consciousness? *Proc Biol Sci*, 279(1737), 2289-2298. doi: 10.1098/rspb.2011.2663
- Milner, A.D., & Goodale, M.A. (2008). Two visual systems re-viewed. *Neuropsychologia*, 46(3), 774-785. doi: 10.1016/j.neuropsychologia.2007.10.005
- Mishkin, M., & Ungerleider, L.G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-occipital cortex in monkeys. *Behav Brain Res*, 6(1), 57-77.
- Mitchinson, Ben, & Prescott, Tony J. (2013). Whisker movements reveal spatial attention: A unified computational model of active sensing control in the rat. *PLoS Comp Biol*, 9(9), e1003236.
- Nelson, Mark E, & MacIver, Malcolm A. (2006). Sensory acquisition in active sensing systems. *Journal of Comparative Physiology A*, 192(6), 573-586.
- Novick, I., & Vaadia, E. (2011). Just do it: Action-dependent learning allows sensory prediction. *PLoS One*, 6(10), e26020. doi: 10.1371/journal.pone.0026020
- O'Regan, J.K., & Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behav Brain Sci*, 24(5), 939-973; discussion 973-1031.
- Okada, K., Rong, F., Venezia, J., Matchin, W., Hsieh, I.H., Saberi, K. (2010). Hierarchical organization of human auditory cortex: Evidence from acoustic invariance in the response to intelligible speech. *Cereb Cortex*, 20(10), 2486-2495. doi: bhp318 [pii] 10.1093/cercor/bhp318
- Ostrovsky, Yuri, Andalman, Aaron, & Sinha, Pawan. (2006). Vision following extended congenital blindness. *Psychological Science*, 17(12), 1009-1014.
- Pascual-Leone, A., & Hamilton, R. (2001). The metamodal organization of the brain. *Prog Brain Res*, 134, 427-445.
- Pasqualotto, A., Lam, J.S., & Proulx, M.J. (2013a). Congenital blindness improves semantic and episodic memory. *Behav Brain Res*, 244, 162-165. doi: 10.1016/j.bbr.2013.02.005
- Pasqualotto, A., & Proulx, M.J. (2012). The role of visual experience for the neural basis of spatial cognition. *Neurosci Biobehav Rev*, 36(4), 1179-1187. doi: 10.1016/j.neubiorev.2012.01.008
- Pasqualotto, A., & Proulx, M.J. (2013). The study of blindness and technology can reveal the mechanisms of three-dimensional navigation. *Behav Brain Sci*, 36(5), 559-560; discussion 571-587. doi: 10.1017/S0140525X13000496
- Pasqualotto, A., Spiller, M.J., Jansari, A.S., & Proulx, M.J. (2013b). Visual experience facilitates allocentric spatial representation. *Behav Brain Res*, 236(1), 175-179. doi: 10.1016/j.bbr.2012.08.042
- Pollok, B., Schnitzler, I., Stoerig, P., Mierdorf, T., & Schnitzler, A. (2005). Image-to-sound conversion: Experience-induced plasticity in auditory cortex of blindfolded adults. *Exp Brain Res*, 167(2), 287-291. doi: 10.1007/s00221-005-0060-8
- Proulx, M.J. (2010). Synthetic synaesthesia and sensory substitution. *Conscious Cogn*, 19(1), 501-503. doi: 10.1016/j.concog.2009.12.005
- Proulx, M.J., Brown, D.J., Pasqualotto, A., & Meijer, P. (2014). Multisensory perceptual learning and sensory substitution. *Neurosci Biobehav Rev*, 41, 16-25. doi: 10.1016/j.neubiorev.2012.11.017
- Proulx, M.J., Stoerig, P., Ludowig, E., & Knoll, I. (2008). Seeing 'where' through the ears: Effects of learning-by-doing and long-term sensory deprivation on localization based on image-to-sound substitution. *PLoS One*, 3(3), e1840. doi: 10.1371/journal.pone.0001840
- Ptito, M., Moesgaard, S.M., Gjedde, A., & Kupers, R. (2005). Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind. *Brain*, 128(Pt 3), 606-614. doi: 10.1093/brain/awh380
- Putzar, Lisa, Gondan, Matthias, & Röder, Brigitte. (2012). Basic multisensory functions can be acquired after congenital visual pattern deprivation in humans. *Developmental neuropsychology*, 37(8), 697-711.
- Qin, W., & Yu, C. (2013). Neural pathways conveying novisual information to the visual cortex. *Neural Plast*, 2013, 864920. doi: 10.1155/2013/864920
- Recanzone, G.H., & Sutter, M.L. (2008). The biological basis of audition. *Annu Rev Psychol*, 59, 119-142. doi: 10.1146/annurev.psych.59.103006.093544
- Reich, L., Szwed, M., Cohen, L., & Amedi, A. (2011). A ventral visual stream reading center independent of visual experience. *Curr Biol*, 21(5), 363-368. doi: 10.1016/j.cub.2011.01.040
- Renier, L.A., Anurova, I., De Volder, A.G., Carlson, S., VanMeter, J., & Rauschecker, J.P. (2010). Preserved functional specialization for spatial processing in the middle occipital gyrus of the early blind. *Neuron*, 68(1), 138-148. doi: 10.1016/j.neuron.2010.09.021
- Renier, L., Collignon, O., Poirier, C., Tranduy, D., Vanlierde, A., Bol, A. (2005a). Cross-modal activation of visual cortex during depth perception using auditory substitution of vision. *Neuroimage*, 26(2), 573-580. doi: 10.1016/j.neuroimage.2005.01.047
- Renier, L., & De Volder, A.G. (2010). Vision substitution and depth perception: Early blind subjects experience visual perspective through their ears. *Disabil Rehabil Assist Technol*, 5(3), 175-183. doi: 10.3109/17483100903253936

- Renier, L., Laloyaux, C., Collignon, O., Tranduy, D., Vanlierde, A., Bruyer, R., & De Volder, A.G. (2005b). The Ponzo illusion with auditory substitution of vision in sighted and early-blind subjects. *Perception*, *34*(7), 857-867.
- Renier, Laurent, Bruyer, Raymond, & Volder, AnneG. (2006). Vertical-horizontal illusion present for sighted but not early blind humans using auditory substitution of vision. *Percept Psychophys*, *68*(4), 535-542. doi: 10.3758/BF03208756
- Renzi, Chiara, Ricciardi, Emiliano, Bonino, Daniela, Handjaras, Giacomo, Vecchi, Tomaso, & Pietrini, Pietro. (2013). The effects of visual control and distance in modulating peripersonal spatial representation. *PLoS One*, *8*(3), e59460.
- Ricciardi, E., Bonino, D., Pellegrini, S., & Pietrini, P. (2014). Mind the blind brain to understand the sighted one! Is there a supramodal cortical functional architecture? *Neuroscience and Biobehavioral Reviews*, *41*, 64-77. doi: DOI 10.1016/j.neubiorev.2013.10.006
- Ricciardi, E., & Pietrini, P. (2011). New light from the dark: What blindness can teach us about brain function. *Curr Opin Neurol*, *24*(4), 357-363. doi: 10.1097/WCO.0b013e328348bdfb
- Ricciardi, Emiliano, Bonino, Daniela, Sani, Lorenzo, Vecchi, Tomaso, Guazzelli, Mario, Haxby, James V. (2009). Do we really need vision? How blind people "see" the actions of others. *The Journal of Neuroscience*, *29*(31), 9719-9724.
- Rockland, K.S., & Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *Int J Psychophysiol*, *50*(1-2), 19-26.
- Röder, Brigitte, Kusmierek, Anna, Spence, Charles, & Schicke, Tobias. (2007). Developmental vision determines the reference frame for the multisensory control of action. *Proceedings of the National Academy of Sciences*, *104*(11), 4753-758.
- Rossetti, Yves, Stelmach, George, Desmurget, Michel, Prablanc, Claude, & Jeannerod, Marc. (1994). The effect of viewing the static hand prior to movement onset on pointing kinematics and variability. *Experimental brain research*, *101*(2), 323-330.
- Sampaio, E., & Dufier, J.L. (1988). [An electronic sensory substitute for young blind children]. *J Fr Ophthalmol*, *11*(2), 161-167.
- Sarlegna, F., Blouin, J., Vercher, J.L., Bresciani, J.P., Bourdin, C., & Gauthier, G.M. (2004). Online control of the direction of rapid reaching movements. *Exp Brain Res*, *157*(4), 468-471. doi: 10.1007/s00221-004-1860-y
- Sathian, K., & Zangaladze, A. (2002). Feeling with the mind's eye: Contribution of visual cortex to tactile perception. *Behav Brain Res*, *135*(1-2), 127-132.
- Schmidt, Robert F. (1981). Somatovisceral sensibility *Fundamentals of sensory physiology* (pp. 81-125): Springer.
- Schroeder, Charles E, Wilson, Donald A, Radman, Thomas, Scharfman, Helen, & Lakatos, Peter. (2010). Dynamics of active sensing and perceptual selection. *Current Opinion in Neurobiology*, *20*(2), 172-176.
- Segond, H., Weiss, D., Kawalec, M., & Sampaio, E. (2013). Perceiving space and optical cues via a visuo-tactile sensory substitution system: A methodological approach for training of blind subjects for navigation. *Perception*, *42*(5), 508-528.
- Segond, H., Weiss, D., & Sampaio, E. (2005). Human spatial navigation via a visuo-tactile sensory substitution system. *Perception*, *34*(10), 1231-1249.
- Soechting, John F. & Flanders, Martha. (1989). Sensorimotor representations for pointing to targets in three-dimensional space. *Journal of Neurophysiology*, *62*(2), 582-594.
- Striem-Amit, E., Dakwar, O., Reich, L., & Amedi, A. (2012). The large-scale organization of "visual" streams emerges without visual experience. *Cereb Cortex*, *22*(7), 1698-1709. doi: 10.1093/cercor/bhr253
- Taira, M., Mine, S., Georgopoulos, A.P., Murata, A., & Sakata, H. (1990). Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Exp Brain Res*, *83*(1), 29-36.
- Thaler, L. (2013). Echolocation may have real-life advantages for blind people: An analysis of survey data. *Front Physiol*, *4*, 98. doi: 10.3389/fphys.2013.00098
- Thaler, L., Arnott, S.R., & Goodale, M.A. (2011). Neural correlates of natural human echolocation in early and late blind echolocation experts. *PLoS One*, *6*(5), e20162. doi: 10.1371/journal.pone.0020162
- Thaler, L., & Goodale, M.A. (2011). Neural substrates of visual spatial coding and visual feedback control for hand movements in allocentric and target-directed tasks. *Front Hum Neurosci*, *5*, 92. doi: 10.3389/fnhum.2011.00092
- Thaler, L., Milne, J.L., Arnott, S.R., Kish, D., & Goodale, M.A. (2014). Neural correlates of motion processing through echolocation, source hearing, and vision in blind echolocation experts and sighted echolocation novices. *J Neurophysiol*, *111*(1), 112-127. doi: 10.1152/jn.00501.2013
- Van Donkelaar, P., Lee, J.H., & Drew, A.S. (2000). Transcranial magnetic stimulation disrupts eye-hand interactions in the posterior parietal cortex. *J Neurophysiol*, *84*(3), 1677-1680.
- Wachowiak, Matt. (2011). All in a sniff: Olfaction as a model for active sensing. *Neuron*, *71*(6), 962-973.
- Ward, J., & Meijer, P. (2010). Visual experiences in the blind induced by an auditory sensory substitution device. *Conscious Cogn*, *19*(1), 492-500. doi: 10.1016/j.concog.2009.10.006
- Weiss, Tali, & Sobel, Noam. (2012). What's primary about primary olfactory cortex? *Nat Neurosci*, *15*(1), 10-12.
- Wolpert, Daniel M, Ghahramani, Zoubin, & Jordan, Michael I. (1995). An internal model for sensorimotor integration. *Science*, *269*(5232), 1880-1882.
- Xu, Jinghong, Yu, Liping, Rowland, Benjamin A, Stanford, Terrence R, & Stein, Barry E. (2012). Incorporating cross-modal statistics in the development and maintenance of multisensory integration. *The Journal of Neuroscience*, *32*(7), 2287-2298.
- Zangaladze, A., Epstein, C.M., Grafton, S.T., & Sathian, K. (1999). Involvement of visual cortex in tactile discrimination of orientation. *Nature*, *401*(6753), 587-590. doi: 10.1038/44139