Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

Review

Multisensory perceptual learning and sensory substitution

Michael J. Proulx [a,*], David J. Brown [b], Achille Pasqualotto [b], Peter Meijer [c]

[a] Department of Psychology, University of Bath, UK
[b] Biological and Experimental Psychology Group, School of Biological and Chemical Sciences, Queen Mary University of London, UK
[c] Metamodal BV, Eindhoven, The Netherlands

ARTICLE INFO

ABSTRACT

One of the most exciting recent findings in neuroscience has been the capacity for neural plasticity in adult humans and animals. Studies of perceptual learning have provided key insights into the mechanisms of neural plasticity and the changes in functional neuroanatomy that it affords. Key questions in this field of research concern how practice of a task leads to specific or general improvement. Although much of this work has been carried out with a focus on a single sensory modality, primarily visual, there is increasing interest in multisensory perceptual learning. Here we will examine how advances in perceptual learning research both inform and can be informed by the development and advancement of sensory substitution devices for blind persons. To allow 'sight' to occur in the absence of visual input through the eyes, visual information can be transformed by a sensory substitution device into a representation that can be processed as sound or touch, and thus give one the potential to 'see' through the ears or tongue. Investigations of auditory, visual and multisensory perceptual learning can have key benefits for the advancement of sensory substitution, and the study of sensory deprivation and sensory substitution likewise will further the understanding of perceptual learning in general and the reverse hierarchy theory in particular. It also has significant importance for the developing understanding of the brain in metamodal terms, where functional brain areas might be best defined by the computations they carry out rather than by their sensory-specific processing role.

Contents

1. Introduction

The human and non-human animal brain undergoes rapid and extensive change during development. A key area of research for neuroscientists concerns the mechanisms of this plasticity from the molecular to the behavioral levels. Many important studies in the last century established that there can be critical periods during development when neuroplasticity is observed (Hubel and Wiesel, 1970). Since that time, however, there has been mounting evidence that even the adult brain retains significant neural plasticity that accompanies perceptual learning (Gilbert et al., 2001).

Studies of perceptual learning have provided key insights into the mechanisms of neuroplasticity and resulting functional neuroanatomy. The central aim of perceptual learning research is to understand how practice of a task leads to either specific or general improvement. Much research on perceptual learning has been fairly low level and unisensory, focusing for example on how

practice results in task-specific improvements in performance and neural plasticity at the level of primary sensory cortex. Of great interest, however, is how generalization can be promoted and some of the most striking evidence for high-level perceptual learning and adult neural plasticity has come from studies of sensory deprivation and sensory substitution devices to overcome such deprivation.

To allow a form of functional 'vision' to occur in the absence of visual input through the eyes, visual information can be transformed by a sensory substitution device into a representation that can be processed as sound or touch, and thus give one the potential to 'see' through the ears or tongue (Bach-y-Rita et al., 1969; Meijer, 1992). Investigations of auditory, visual and multisensory perceptual learning can have key benefits for the advancement of sensory substitution, and the study of sensory deprivation and sensory substitution likewise will further the understanding of perceptual learning.

Although there have been numerous studies examining visual, auditory, and multisensory perceptual learning over the past 50 years (Gibson, 1963; Goldstone, 1998), there has not been a synthesis that brings these findings together under the same theoretical structure. Here we bring together advances on the reverse hierarchy theory of perceptual learning (Ahissar and Hochstein, 2004) and the metamodal hypothesis of brain organization (Pascual-Leone and Hamilton, 2001) to provide a behavioral and neural explanation of visual, auditory, and multisensory perceptual learning (Ghazanfar and Schroeder, 2006; Shams and Seitz, 2008). Certainly some aspects are better understood at a behavioral level, and yet other aspects at a neural level, and this synthesis of the reverse hierarchy and metamodal theories will highlight areas where such cross-fertilization of research efforts would be beneficial and specify possible constraints for each theory. We also provide an examination of the reciprocal benefits of sensory deprivation and sensory substitution devices as means to understand the mechanisms and neural basis of perceptual learning. This approach will likely also provide further advances for the development of sensory substitution to aid those with sensory impairments.

2. Visual perceptual learning and the reverse hierarchy theory

Psychophysical studies of visual perceptual learning have established that practicing a task results in improvement that is often restricted to the stimuli used during training (Fiorentini and Berardi, 1980; McKee and Westheimer, 1978). The specificity of improved performance is taken to indicate that neural plasticity manifests at the 'low' level of primary visual cortex because the neurons at that level have receptive field properties for the particular visual features that have been learned. This use of psychophysical findings to constrain the possible neural basis of perceptual learning was termed 'psycho-anatomy' by Julesz (1972).

Training studies have demonstrated the specific improvement of performance for a number of visual features that are often spatial in nature, such as vernier acuity (Beard et al., 1995; Fahle et al., 1995; McKee and Westheimer, 1978; Poggio, 1995; Saarinen and Levi, 1995), orientation and texture (Karni and Sagi, 1991; Vogels and Orban, 1985), motion (Ball and Sekuler, 1982, 1987), and spatial frequency (Fiorentini and Berardi, 1980, 1981). What sort of specificity is normally reported? Learning can be spatially specific such that training in one visual field does not transfer to another (Karni and Sagi, 1991). It can also be feature specific, such that training with one orientation does not transfer to another orientation (Karni and Sagi, 1991). It is important that the underlying mechanisms of such specific perceptual learning have often been described as the retuning of low level sensory areas in the brain. Psychophysical

experiments and modeling by Dosher and colleagues have demonstrated that such neural plasticity can, however, be accomplished in other ways such as the reweighting of the visual channels used for a task (Dosher and Lu, 1998; Petrov et al., 2005).

There have been some surprising cases of generalization, however, that seemed to contradict the findings of specific perceptual learning. For example, although a previous report found training benefits restricted to one region of space, and even one eye (Karni and Sagi, 1991), a subsequent study found that a similar texture discrimination task could transfer from one eye to the other (Schoups et al., 1995). The 'reverse hierarchy theory' of visual perceptual learning (Ahissar and Hochstein, 2004) was developed to account for apparently conflicting findings such as this. The reverse hierarchy theory posits that the difficulty and characteristics of a task determine the level of cortical processing at which attentional mechanisms are required (see Fig. 1). An easier task that can be carried out on the basis of more general levels of feature discrimination instead drive processing and attentional resources to higher level cortical association areas, such as the lateral intraparietal area with its larger receptive fields. The harder the task and the more specific the discrimination required, the more it tends to drive processing and attentional resources to lower, primary sensory areas, such as V1 with its smaller receptive fields. The idea is that perceptual learning can occur at all cortical levels of processing: initially higher-level areas would be recruited, however feedback connections to lower-level areas would be employed if necessary. When perceptual learning occurs at higher-level areas, then the training can generalize to other regions of space and to other features (Pavlovskaya and Hochstein, 2011). However, when perceptual learning occurs at lower-level areas, then the training will remain specific to the spatial locations and features used during training. Importantly the use of feedback connections for perceptual learning has gained support from findings in vision (Juan et al., 2004; Zhang et al., 2008), and in audition (Wong et al., 2007).

The role of higher-level cortical areas, rather than just lower-level striate (V1) and extrastriate areas, in perceptual learning, as proposed by reverse hierarchy theory, has been confirmed by other findings in the literature. For example, a novel paradigm that involved double training, where one retinal location was exposed to the relevant task and another to an irrelevant task (Xiao et al., 2008). A transfer of perceptual learning was induced by the irrelevant training at the second location, suggesting that higher order, nonretinotopic brain areas were involved in learning and thus promoted location generalization (Dosher and Lu, 1998; Petrov et al., 2005; Wang et al., 2012; Zhang et al., 2010). Reverse hierarchy theory has provided a framework to characterize both the specific and generalized perceptual learning in vision, and the recruitment of cortical areas along the hierarchy of visual information processing.
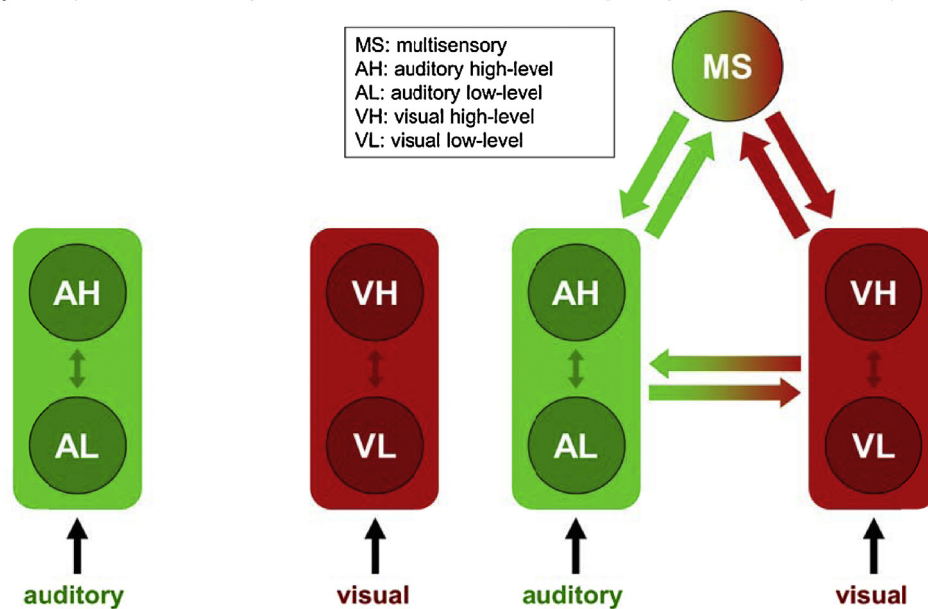
3. Auditory perceptual learning

Compared to the abundance of literature on visual perceptual learning, the literature on the specificity and transfer of auditory perceptual learning is scarce, though with Wright and colleagues making many of the seminal contributions to this field in recent years (Wright and Zhang, 2009). As with the vision literature, the first aim in auditory research was to establish whether practice improves performance on auditory tasks. The primary features of interest in the auditory domain are frequency (spectral) information and temporal information, such as the order, interval or duration of stimuli. These features are particular important for speech and music perception in humans. Moreover, temporal cues can be important for spatial localization as well (Jeffress, 1948).

A prototypical paradigm for the study of auditory perceptual learning is a temporal discrimination task (Wright et al., 1997).

A. Unisensory Reverse Hierarchy Theory
learning occurs at low (specific) and/or high
(general) levels depending on the task

B. Multisensory Reverse Hierarchy Theory
adds learning with multisensory reinforcement
and learning along multisensory pathways

MS: multisensory
AH: auditory high-level
AL: auditory low-level
VH: visual high-level
VL: visual low-level

Fig. 1. Figure depicting a unisensory and a multisensory reverse hierarchy theory of perceptual learning. (A) Unisensory learning is shown as being modality specific, such that an auditory task (green) is supported by either low-level auditory areas for specific learning or high-level auditory areas for general learning (Ahissar, 2001). A visual task (red) exhibits similar activity in visual areas (Ahissar and Hochstein, 2004), again with low- and high-level areas defined in afferent terms, with low-level corresponding to primary sensory areas and high-level to association cortex. (B) Multisensory learning is shown to represent the same possible mechanisms for two different conditions: first, learning under multisensory stimulation can lead to correlated activity in higher-level multisensory areas (Shams and Seitz, 2008); second, learning can progress from primary sensory areas to higher-level multisensory areas under complex unisensory stimulation. Activity may then cascade back down the hierarchy such that generalization across modalities occurs when these higher-level, multisensory areas are implicated in learning either unisensory or multisensory tasks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

A participant is presented with two stimuli. One is denoted the standard and consists of two tones separated by a standard time interval, say 100 ms. The comparison stimulus has a small amount of time added to the standard interval, say 10 ms, to assess whether the participant can discriminate between the two. Several studies have reported that training leads to improved performance on this specific task (Fitzgerald and Wright, 2005; Wright et al., 1997).

This paradigm can also be used to test for generalization of learning by changing the frequency of the tones to assess spectral generalization, and by changing the standard temporal interval (to be shorter or longer) to assess temporal generalization. Thus far there is evidence in the literature of generalization to new frequencies, but not to new temporal intervals, with training (Fitzgerald and Wright, 2005; Wright et al., 1997). This depends on the nature of the task, however. Another study examined the learning of interaural time and intensity differences; these are features that are important for spatial localization (Wright and Fitzgerald, 2001). In this study of time and intensity differences, participants did not transfer learning to untrained frequencies. Although the majority of studies have not found much generalization of auditory learning in the temporal tasks (Karmarkar and Buonomano, 2003; Wright et al., 1997), one study of duration discrimination did find generalization to an untrained temporal interval (Lapid et al., 2009).

Similar studies have addressed auditory perceptual learning for frequency discrimination tasks. Humans can improve in the specific discrimination of a particular trained frequency, with some generalization to untrained frequencies, though with a gradient of poorer performance (Irvine et al., 2000). Similarly, another study reported that learning curves during a second phase of training were steeper for previously untrained frequencies than for trained ones (Demany

and Semal, 2002). Still other studies have looked into the effect of training on discrimination of changes in complex spectra (Gockel and Colonius, 1997; Green, 1983).

In contrast to the lack of generalization observed in temporal discrimination tasks, frequency discrimination learning has been found to partially generalize to untrained stimulus durations and intervals, and across ears (Delhommeau et al., 2002). For example, complete generalization is possible as demonstrated by similar improvement from the trained to untrained ear (Micheyl et al., 2006a,b). There is also evidence that learning on frequency discrimination generalizes across conditions in which the pure tone frequency is fixed or 'roves' across frequencies. Both wide and narrow frequency bands generalize to the fixed frequency with generalization from the fixed to the narrow band in poor listeners (Amitay et al., 2005).

When considering frequency and amplitude, training on pure tones was found to generalize to complex tones containing harmonics of the fundamental frequency that can be resolved by the peripheral auditory system but not to stimuli with unresolved harmonics (Demany and Semal, 2002; Grimault et al., 2003). In addition there was no generalization to noise bands modulated at the fundamental frequency implying that the auditory system uses two processes to encode pitch and is dependent on the resolution of harmonics. The amount of learning was found to be greater when the harmonics could be processed by the low-level peripheral auditory system.

A further intriguing result from frequency discrimination was shown by Amitay et al. (2005) who found an improvement in discrimination thresholds even when the trained stimuli were classed as impossible to discriminate between, that is when the trained frequency differentials were 0 Hz. That improvement in this impossible condition compared favorably to easy (400 Hz) and

difficult (7 Hz) conditions led the authors to posit that training, rather than adapting a stimulus comparison mechanism, may just improve the ability to attend and access low-level representations of the stimulus that are task specific. A further explanation for these counterintuitive results was proposed by Micheyl et al. (2006a,b) who suggested that the random variability of neural responses to auditory stimuli may render the different representations of identical stimuli as qualitatively different, and this may be sufficient to fine tune the relevant stimulus comparison process necessary for perceptual learning. Although the literature on auditory perceptual learning is smaller than that of visual learning, the specific or generalized transfer of such learning is beginning to be understood for a variety of auditory features and tasks.

4. Multisensory and cross-modal perceptual learning and generalization

Although audition is clearly important for tasks such as speech perception, the influence of other sensory inputs, particularly vision, should not be underestimated (Shams and Seitz, 2008). Given the scope of this review, here we will focus on multisensory processes more generally, rather than the more specific phenomenon of multisensory integration (Stein et al., 2010). For example, the integration of visual input with its associated auditory signal, even if arbitrary, allows for improved discrimination of the target signal from the background noise (Seitz et al., 2006). This is exemplified in Cherry's (1953) 'cocktail party' situation where isolation of the target's voice is dependent on the integration of the specific auditory signal with the relevant facial movements. The influence of vision on speech perception was demonstrated memorably by the McGurk effect in which lip movements incongruent to the spoken phoneme elicit a perceived phoneme that differs from the actual utterance. For example, the auditory 'ba' overdubbed onto the visual production of 'ga' is perceived as 'da' (McGurk and MacDonald, 1976).

Numerous behavioral studies have shown that synchronous presentation of audiovisual stimuli facilitates superior speech recognition in comparison to its unimodal counterpart (Ross et al., 2007; Seitz et al., 2006; Shams and Seitz, 2008; Sumby and Pollack, 1954), and that the more reliable modality has the greater influence on the eventual perception (Alais and Burr, 2004). Interestingly, the neural basis for audiovisual integration in speech perception appears to rely in part on what have been considered unimodal cortical areas. In response to auditory speech, bilateral activation of the superior temporal gyrus and anterior and posterior regions of the superior temporal sulcus have been reported, while left frontal regions such as Broca's area are involved in phonetic recoding and auditory discrimination tasks (Zatorre et al., 1992). For visual speech processing such as lip-reading, activation is found in early visual areas in the occipital cortex with further processing in the posterior region of the STS, inferior frontal gyrus and premotor cortex (Olson et al., 2002; Scott and Johnsrude, 2003). Importantly, there are reports that phonetic perceptual learning exhibits not only specific perceptual learning, but generalization as well in both infant and adult participants (Hervais-Adelman et al., 2011; Kraljic and Samuel, 2006; Maye et al., 2008). It is also interesting to note that sleep has been found to be an important enabler of promoting generalization of phonological categories (Fenn et al., 2003). Across studies, the superior temporal sulcus appears the likely candidate for the cross-modal integration in audiovisual speech perception as not only does it show activation in both auditory and visual speech processing but it also receives convergent inputs from both the auditory and visual cortices (Powers et al., 2012). Indeed, the adjoining superior temporal gyrus appears to even encode species-specific multisensory information, such as using human voice sources to judge source size information (von Kriegstein et al., 2007). Transcranial magnetic stimulation has revealed that the superior temporal sulcus plays a pivotal role in auditory–visual integration, such as that required for the McGurk effect (Beauchamp et al., 2010).

There are many fascinating demonstrations of cross-modal influences and mappings across sensory modalities in perception and cognition in domains other than language. The ubiquity of such cross-modal influences suggests that perceptual learning might extend across modalities as well. Sometimes these effects are unidirectional. For example, Grahn et al. (2011) reported that an internal auditory rhythm representation may be triggered during visual rhythm perception, though not vice versa. In other cases, the cross-modal influence goes in both directions. Kim et al. (2010) found an auditory influence on a visual motion task. In the other direction, King (2009) found visual influences on auditory spatial learning whilst Dellorso and Schirillo (2010) found similar visual influences on multiple internal auditory stimuli. The cross-modal influences extend to the chemical sensory modalities as well. For example, Simner and Haywood (2009) presented evidence that there are cross-modal mappings between gustation and audition. Perceptual learning can also generalize across sensory modalities. Nagarajan et al. (1998) showed that while training-related improvements in somatosensory interval discrimination were temporally specific within the tactile modality there was generalization not only within this modality to skin location, but to the auditory modality as well.

Transfer to other modalities does not occur with all tasks, however. When evaluating whether training on auditory duration discrimination facilitates visual duration discrimination, Lapid et al. (2009) found no evidence of a cross-modal transfer effect from the auditory to the visual modality for this task. Yet, the vast majority of the studies on cross-modal perceptual learning have reported some form of cross-modal transfer across modalities. Planetta and Servos (2008), using a tactile duration discrimination task, found generalization to motor interval production with the same temporal specifications. Sometimes the temporal intervals can transfer in a more generalized fashion, such as in a study by Bartolo and Merchant (2009), who used an auditory interval reproduction task and found generalization to vision.

Given these complexities, when does perceptual learning generalize across sensory modalities? First, it appears that to have perceptual learning with cross-modal generalization, stimuli have to share some spatiotemporal attributes (i.e., the same duration or same location) or possess some characteristics that facilitate the transfer (e.g., certain durations are more likely to be generalized, see Bartolo and Merchant, 2009). Such learning does not have to be task relevant or at the level of awareness. As long as there is some implicit correlation between stimuli, such as in space and time, there is the potential for task irrelevant perceptual learning as demonstrated in a number of studies by Seitz and Watanabe (Seitz and Watanabe, 2009; Tsushima et al., 2008). One clear example of such generalized learning to another, seemingly task irrelevant modality can be observed by the improved tactile spatial acuity in pianists (Ragert et al., 2004). Second, the characteristics that are more salient for a given task are more likely to be generalized across modalities (Jain et al., 2010). A combination of psychophysical methods to limit the psycho-anatomical basis for such generalization (Julesz, 1972) and neurophysiological studies to limit the neural spatiotemporal correlates of learning (Bruns et al., 2011; van der Linden et al., 2011) will help to predict when such generalization occurs, given that spatiotemporal congruence appears to be at the core of multisensory integration and processing (Lewald et al., 2001; Macaluso and Driver, 2005; Meredith and Stein, 1986).

Standard Functional Architecture and the Metamodal Hypothesis

Auditory   **Visual**
Temporal   *Spatial*
Information *Information*
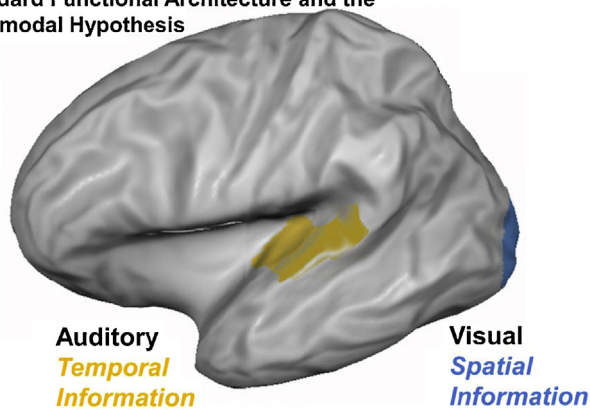
Fig. 2. Figure depicting a simplified version of the metamodal brain hypothesis. The primary sensory cortices are labeled in terms of the form of information processed. In blue, primary visual cortex processes spatial information. In yellow, primary auditory cortex processes temporal information (Pascual-Leone and Hamilton, 2001). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

## 5. Sensory substitution, blindness and the metamodal brain

Generalization of learning has also been demonstrated in studies using sensory substitution devices developed for translating images into formats that the blind or blind-folded can process with sensory modalities other than vision. Kim and Zatorre (2008) incorporated trained and novel stimuli, of varying complexity, into a test phase to examine the breadth of generalization possible in visual-to-auditory substitution. Their findings indicated that visual-to-auditory generalization to novel stimuli occurred within the first week of training and was significantly influenced by stimulus complexity. Comparable results were found in a later study examining the auditory substitution of shape using tactile feedback (Kim and Zatorre, 2010). This study also demonstrated that abstraction from stimuli learned in the trained modalities (tactile-auditory) would transfer to an untrained modality (vision).

This result supports the idea of a metamodal (Pascual-Leone and Hamilton, 2001) or supramodal (Kupers et al., 2011; Kupers and Ptito, 2011; Ricciardi and Pietrini, 2011) organization of the brain, depicted in Fig. 2, where brain areas are involved in a particular form of information processing (e.g., shape recognition) independently from the sensory modality that produced the input. Research in sighted individuals has found that different brain regions are responsive to particular forms of visual stimuli. Two well-known examples are the fusiform face area (Kanwisher et al., 1997) and the parahippocampal place area (Epstein and Kanwisher, 1998) that respond preferentially to images of faces and places, such as pictures of houses, respectively. The metamodal hypothesis was developed to account for evidence that the same brain regions normally responsive to visual categories maintain such response selectivity in the absence of visual experience (Ricciardi and Pietrini, 2011). For example, the lateral occipital area is involved in the perception of object shape and identity. This was discovered originally with visual stimuli in sighted individuals. A study of congenitally blind participants using tactile stimuli (such as bottles and shoes) found that this same region of extrastriate cortex was still activated for shape and object discrimination, and thus must be supramodal due to its modality-independent role in shape perception (Pietrini et al., 2004). Furthermore the same study found that regions of inferior temporal cortex responded similarly for either the visual or tactile recognition of objects in sighted participants as well, thus also suggesting that area of cortex is supramodal. These results have also been supported by converging evidence in

the auditory domain by using sensory substitution. A follow-up study by Kim and Zatorre (2011) showed that an area once thought to be visual, the lateral occipital complex, an area involved in shape recognition (Amedi et al., 2002), was active during the shape recognition task carried out with auditory stimulation using the sensory substitution device. These studies again provided validation of the metamodal hypothesis for the functional organization of the brain, as have several other studies reviewed by Ricciardi and Pietrini (2011). Interestingly, the sensory-neutral representation of shape has been validated with sensory substation devices that specifically use stimulation to sensory organs that are not normally used for the purpose of shape recognition or spatial navigation, such as hearing with the ears (Amedi et al., 2007; Brown et al., 2011; Proulx et al., 2008) and feeling with the tongue (Chebat et al., 2011; Matteau et al., 2010; Proulx and Stoerig, 2006; Ptito et al., 2005).

A metamodal organization of the brain is not the only possible explanation for these results, however. First there is a distinction between having either supra- or metamodal representations and cross-modal plasticity (Kupers et al., 2011; Ricciardi and Pietrini, 2011). An extreme interpretation would be that metamodal representations exist independent of developmental experience and that cross-modal plasticity would come about through a rewiring of cortical connections in sensory deprivation (Collignon et al., 2011; Frasnelli et al., 2011; Kupers et al., 2011; Kupers and Ptito, 2011; Noppeney, 2007; Ricciardi and Pietrini, 2011). However there has not yet been any evidence of brand new neural connections, but rather only modification and strengthening of existing connections (Ptito et al., 2008), and increased functional connectivity (Kupers et al., 2006; Ptito et al., 2005). Therefore the apparent recruitment of visual cortex, for example, in auditory or tactile processing in blindness or visual deprivation arises from the use of existing neural connections. In this sense, cross-modal plasticity does not necessarily provide a competing hypothesis for the metamodal organization of the brain, but rather is a mechanism of perceptual learning that unmasks existing connections and consolidates and intensifies them (Kupers et al., 2011).

The hypothesis of a metamodal brain is corroborated by the effects produced by long-term blindfolding of sighted participants, for example with three weeks of sensory deprivation along with visual-to-auditory sensory substitution training (Pollok et al., 2005; Proulx et al., 2008). This resulted in effects of neuroplasticity in the form of brain activation equivalent to blind participants in purportedly visual cortical areas but with auditory stimulation (Boroojerdi et al., 2000; Facchini and Aglioti, 2003; Lewald, 2007; Merabet et al., 2008). Such effects can occur in a matter of days rather than weeks. For example, changes in functional brain activation have been observed with only five days of visual deprivation and Braille training for sighted participants (Kauffman et al., 2002; Pascual-Leone and Hamilton, 2001). However, the rapid changes in adult participants are not compatible with the establishment of novel neural connections, which therefore must already be in place. Such connections certainly do exist naturally and the primary sensory areas of the primate brain have been found to have afferent connections from multiple sensory modalities (Falchier et al., 2002; Murata et al., 1965; Rockland and Ojima, 2003). The observed rapid changes may therefore be accounted for by unmasking of existing neural connections (Pascual-Leone and Hamilton, 2001). Moreover, recent studies reported that sensory loss does not change the functional specialization of a given cortical region and instead just changes the origin of the input; see Meredith et al. (2011) for deafness and Renier et al. (2010) for blindness. In addition, a number of reviews of studies on humans have reported that the visual cortex is functionally involved in tactile tasks also in sighted subjects, and these provide additional evidence for the metamodal nature of the 'visual' brain areas revealed by tactile stimulation (Driver and Noesselt, 2008; Ghazanfar and Schroeder, 2006; Sathian and Zangaladze,
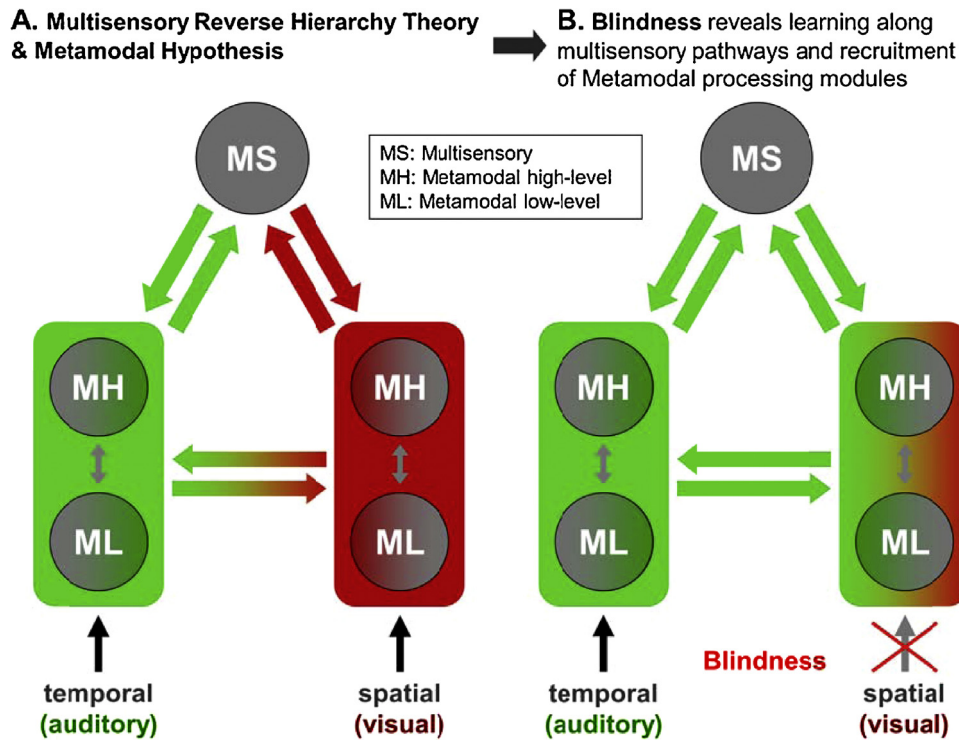
Fig. 3. A depiction of the implications of metamodal brain organization for perceptual learning. (A) Similar to Fig. 1B, perceptual learning is here shown in a reverse hierarchy framework, however visual information (red) is conceptualized as the primary modality for spatial tasks, and auditory information (green) is conceptualized as the primary modality for temporal tasks (Pascual-Leone and Hamilton, 2001). (B) The impact of blindness on perceptual learning is illustrated by removing the visual (red) input. In the metamodal hypothesis, so-called visual cortex is the area responsible for spatial processing. Thus if presented with a task that would require the spatial abilities normally attributed to vision, such as shape perception (Amedi et al., 2007), then auditory stimulation (green) can activate and induce perceptual learning in these 'visual' areas. Note that the visual areas might be able to maintain visual processing such that cross-modal activation of these areas can give rise to visual phenomenology as reported by Ward and Meijer (2010). See also Fig. 2 for a simple depiction of the neuroanatomy of the metamodal hypothesis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

2002; Zangaladze et al., 1999). Finally, Poirier et al. (2006) reported that blind participants exhibited equivalent patterns of activation of visual and auditory areas as sighted participants during auditory motion judgments.

The areas of the brain appear to be organized by sensory modality. However, this might instead reveal which sensory modality is dominant for a particular form of computation, such as relying on retinotopic maps in visual cortex for spatial tasks (Pasqualotto and Proulx, 2012). As noted by Pascual-Leone and Hamilton (2001), it matters less where information originates, but instead it matters more what the information is for so one can respond appropriately. This is sensible considering that electrochemical information transfer is the basis of the neural code, so there are no a priori restrictions on information processing due to its origin or sensory modality.

6. A metamodal framework for multisensory perceptual learning

The perceptual learning evidence described in previous sections can be mapped on to a metamodal framework, depicted here in Fig. 3. The key idea is that auditory brain areas contain the processing modules that are best at detecting and discriminating temporal patterns. Although there is evidence for the representation and segregation of both what and where information in both auditory and visual cortex (Collignon et al., 2011), particularly with sensory substitution, the auditory cortex is still optimal for the processing of temporal information. For example, Bao et al. (2004) reported non-auditory temporal discrimination in auditory areas of the brain. Likewise the visual cortical areas might contain the processing modules that are best at detecting and discriminating

spatial patterns. Thus when Grahn et al. (2011) found that visual rhythm perception evokes an internal auditory rhythm representation, this is due to the temporal nature of the task. In contrast, a spatial task, even with auditory presentation, will evoke visual representations (King, 2009). Figs. 2 and 3 show how the sensory cortices can be classified by the form of information they process best: spatial processing in visual cortex, and temporal processing in auditory cortex. Although these are each sensory-specific areas, they also can receive information either directly or indirectly from other sensory modalities (Pascual-Leone and Hamilton, 2001).

7. A reverse hierarchy theory of multisensory perceptual learning

The metamodal theory of functional brain organization provides an explanation of findings about the physical basis of neural information processing. Theories of perceptual learning, such as the reverse hierarchy theory, provide an explanation of the cognitive basis for information processing. The union of these two theories can therefore provide a more complete view of the literature on perceptual learning and provide additional explanatory power and predictions that are not possible with either theory alone. A unified model for multisensory perceptual learning should describe brain areas not by the source of their inputs (i.e., eyes, ears, etc.), but instead by activation determined by the cognitive level and physical type of processing required by the task, as depicted in Fig. 3.

The transfer of perceptual knowledge across modalities is a complex adaptive phenomenon that takes into account the spatiotemporal congruence of the stimuli and their relevance for the task (Angelaki et al., 2009; Deneve and Pouget, 2004; Holmes and

Spence, 2005). This can explain the somewhat mixed results associated with the transfer of perceptual knowledge across modalities. In their review on multisensory learning, Shams and Seitz (2008) support the idea that multisensory learning improves learning performance because multisensory information produces broader brain activation involving both unisensory and multisensory areas (see Fig. 1B). Additionally, they note that human cognition evolved in a multisensory environment while our daily interaction with the external world is multisensory too, such that multisensory learning is more ecologically relevant than unisensory perceptual learning.

Must the multisensory model of perceptual learning operate independently of the standard unisensory model? The primary difference between these two states of perceptual learning is whether there is unisensory stimulation or correlated multisensory stimulation. This is illustrated in Fig. 1A where learning progresses from primary sensory areas to higher-level areas for that modality, under unisensory stimulation, or instead in Fig. 1B where learning under multisensory stimulation leads to correlated activity in higher-level multisensory areas. As a result of the differential stimulation, cortical areas related either to one sensory modality or to multisensory processing are implicated in learning that might then be either specific or generalized. Reconsidering this from the perspectives of the metamodal brain theory and the reverse hierarchy theory can provide a unified description of perceptual learning and make novel predictions for future work.

First consider how the metamodal perspective impacts both the unisensory and multisensory states, as depicted in Fig. 3. Research on sensory deprivation (Amedi et al., 2007; Merabet et al., 2005; Pasqualotto and Proulx, 2012) has revealed that a number of spatial tasks activate visual cortical regions, such as Braille reading, auditory tasks, and allocentric spatial tasks. Although the input could be auditory, and certainly activate primary auditory cortex, the spatial nature of a task like shape perception would also activate 'visual' cortex to carry out the necessary computations (Amedi et al., 2007). Merabet et al. (2009) found that disruption of visual cortex processing through TMS impaired the ability of a blind user of a visual-to-auditory sensory substitution device to identify objects, thus making the functional relevance of the so-called visual cortex essential to processing in what otherwise would have been considered to be an auditory task. This occurs in the tactile domain as well. The functional relevance of the activation of visual areas in haptic tasks was corroborated by TMS studies that transiently blocked the activity of the visual cortex producing a drop in the haptic task performance (Cohen et al., 1997; Sathian, 2005). At this time it is unclear whether these cross-modal interactions might occur through direct connections between primary sensory areas or through processing routed through multisensory areas in association cortex; both possibilities are illustrated in Fig. 3.

The key to unifying unisensory and multisensory perceptual learning with the metamodal perspective appears to be stimulus or task complexity, which is a core aspect of the reverse hierarchy theory. Compared to the unisensory standard model of perceptual learning, tasks that are either multisensory in nature or require complex learning are richer and less likely to be defined by single low-level features. At the very least a conjunction of several features or stimuli defines this richness, such as congruent audiovisual stimulation or the complex spatiotemporal patterns in Braille reading. Although Braille reading has been found to stimulate and even require low-level visual cortex in the blind (Cohen et al., 1997) or blindfolded (Kauffman et al., 2002; Pascual-Leone and Hamilton, 2001), clearly this must be mediated by higher-order areas that, at the very least, re-route such processing from primary somatosensory cortex to visual cortex for spatial decomposition, recognition, and reading. As a result, learning takes place in higher-order areas beyond a single primary sensory cortical area, thus making generalized learning possible beyond the specific learning gains that would

otherwise be expected. This is depicted in Fig. 3A, where learning might progress from primary sensory areas to higher-level areas with multisensory input and then follow feedback connections to yet other primary sensory areas for another modality. Candidate multisensory areas could be high-level unisensory areas that have revealed multisensory input such as anterior inferotemporal cortex (Cappe et al., 2009; Iwai et al., 1987; Tompa and Sary, 2010), or even multisensory areas, such as the superior temporal gyrus (Shams and Seitz, 2008) or posterior superior temporal sulcus (Powers et al., 2012). The combination of larger receptive fields and weak feature preferences make these multisensory or higher level areas likely candidates for the neural basis of cross-modal influences reported in the literature (Grahn et al., 2011; King, 2009). In addition, studies on multisensory processing and integration (Beauchamp, 2005; Calvert and Thesen, 2004; Small, 2004) suggest that associative areas should not be too specific for pairs or triplets of sensory modalities, again suggesting these areas are good candidate areas for consideration in future research given that any areas with multisensory processing responses, even high level areas once thought to be unisensory, are also truly metamodal in nature. It might be that some multisensory areas are restricted to particular combinations of modalities, such as visual–auditory or visual–tactile, however moving further up the hierarchy would likely lead to omnimodal areas responsive to all sensory input. Clearly further research delimiting a functional hierarchy of sensory modality processing and interactions would be an exciting avenue for investigation.

A multisensory reverse hierarchy theory could therefore be used to derive new predictions. One such example would be that a unisensory stimulus might be able to draw on cortical processing in higher order areas that are multisensory in nature (see Figs. 1B and 3A). Although research in auditory perceptual learning of temporal interval discrimination, for example, seems to generalize to novel frequencies after a long bout of learning, it does not generalize to novel temporal intervals (Wright et al., 2010). A new prediction could be that a richer stimulus made up of a band of frequencies would drive perceptual learning to a higher order level of cortical processing that would then also make generalization to new temporal intervals more likely. Ahissar et al. (2009) noted that, for an auditory reverse hierarchy theory, if perceptual learning is possible with a more complex stimulus, then specific benefits could be seen at both lower and higher level representations. Lower levels of cortical processing would be sufficient for improvements with specific frequency bands and temporal resolution (Batra et al., 1997), but higher levels of cortical processing that integrate across spectral and temporal categories would be necessary for generalization (Zatorre et al., 2004).

Further predictions could be made regarding the nature and extent of cross-modal transfer of learning. In particular the merging of reverse hierarchy theory with the metamodal theory of brain organization suggests that cross-modal transfer should occur in situations where the task relies on computations that are normally associated with a primary sensory area other than that which is being stimulated. Therefore, on the basis of findings in visual perceptual learning (Ahissar and Hochstein, 1997), complex multisensory tasks might shift neural activation to areas performing both high and low level information processing, while a less demanding multisensory task will mainly activate the areas performing high level processing (Pavlovskaya and Hochstein, 2011). Thus, in the case of a complex task the 'extra' activation will promote strong cross-modal generalization. Findings that learning in a somatosensory interval discrimination task generalized to the auditory modality suggest that perhaps the superior temporal processing in the auditory modality is used as the metamodal area for such temporal discrimination tasks (Nagarajan et al., 1998). Ultimately it should be possible to distinguish the role of the richness

of the input from the nature of the task to determine whether it is the reverse hierarchy or the metamodal aspect that mostly drives the specific or general level of learning that is found.

8. Conclusion

A greater understanding of how complex unisensory stimulation might promote cross-modal generalization is interesting in its own right, but also for understanding the basis of sensory substitution. A key goal of sensory substitution work is to provide the full suite of cognition and behavior that normally accompanies the missing sensory modality, including the phenomenal experience of it (Proulx, 2010; Ward and Meijer, 2010). Numerous studies have reported activity in putative visual areas of the brain during auditory or tactile stimulation by a sensory substitution device (Amedi et al., 2007; De Volder et al., 1999; Ptito et al., 2005; Renier et al., 2005). Although some long-term users of such devices have reported great facility at using the device and even phenomenal 'visual' experience (Ward and Meijer, 2010), the difficulty of learning to use such a device seems to inhibit widespread adoption of their use. Key to such a device is a form of cross-modal transfer of perceptual learning, where one can receive visual information via auditory stimulation from a device like The vOICe (Meijer, 1992) and associate that with tactile feedback. A greater understanding of multisensory perceptual learning can arise from the use of sensory substitution as an experimental tool to examine such cross-modal learning and generalization. Furthermore, a greater understanding of a metamodal reverse hierarchy theory would allow for advances in sensory substitution that best take advantage of how the human brain achieves multisensory perceptual learning.

Although the primary focus here was to unify research in the visual and auditory modalities, further theoretical development should be able to extend this account to other modalities only briefly mentioned here, such as tactile perception. There are certainly a number of studies that have examined specific perceptual learning, generalization, and the impact of sensory deprivation in the tactile domain (Karim et al., 2006; Sathian and Zangaladze, 1998; Wong et al., 2011), making this a theoretically rich area to explore.

Perceptual learning can offer insights in how neural plasticity affects brain activity and human behavior. In particular, here we aimed to integrate the reverse hierarchy theory, which offers an elegant paradigm to interpret empirical results and forge new predictions, with the metamodal brain theory, which provides a novel paradigm to interpret brain functioning and development. The integration of these two perspectives enables a more compelling interpretation of the often counterintuitive results from studies on sensory deprivation and sensory substitution, which represent the vanguard of the research on neural plasticity. Future investigations will need to assess whether multisensory perceptual learning could mitigate the impact of sensory deprivation (e.g. blindness), and whether it could improve the learning of novel skills.

Acknowledgements

References

Ahissar, M., 2001. Perceptual training: a tool for both modifying the brain and exploring it. Proceedings of the National Academy of Sciences of the United States of America 98 (21), 11842–11843, http://dx.doi.org/10.1073/pnas.221461598.

Ahissar, M., Hochstein, S., 1997. Task difficulty and the specificity of perceptual learning. Nature 387 (6631), 401–406, http://dx.doi.org/10.1038/387401a0.

Ahissar, M., Hochstein, S., 2004. The reverse hierarchy theory of visual perceptual learning. Trends in Cognitive Sciences 8 (10), 457–464, http://dx.doi.org/10.1016/j.tics.2004.08.011.

Ahissar, M., Nahum, M., Nelken, I., Hochstein, S., 2009. Reverse hierarchies and sensory learning. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 364 (1515), 285–299, http://dx.doi.org/10.1098/rstb.2008.0253.

Alais, D., Burr, D., 2004. The ventriloquist effect results from near-optimal bimodal integration. Current Biology 14 (3), 257–262, http://dx.doi.org/10.1016/j.cub.2004.01.029.

Amedi, A., Jacobson, G., Hendler, T., Malach, R., Zohary, E., 2002. Convergence of visual and tactile shape processing in the human lateral occipital complex. Cerebral Cortex 12 (11), 1202–1212.

Amedi, A., Stern, W.M., Camprodon, J.A., Bermpohl, F., Merabet, L., Rotman, S., Hemond, C., Meijer, P., Pascual-Leone, A., 2007. Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. Nature Neuroscience 10 (6), 687–689, http://dx.doi.org/10.1038/nn1912.

Amitay, S., Hawkey, D.J., Moore, D.R., 2005. Auditory frequency discrimination learning is affected by stimulus variability. Perception & Psychophysics 67 (4), 691–698.

Angelaki, D.E., Gu, Y., DeAngelis, G.C., 2009. Multisensory integration: psychophysics, neurophysiology, and computation. Current Opinion in Neurobiology 19 (4), 452–458, http://dx.doi.org/10.1016/j.conb.2009.06.008.

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.

Ball, K., Sekuler, R., 1982. A specific and enduring improvement in visual motion discrimination. Science 218 (4573), 697–698.

Ball, K., Sekuler, R., 1987. Direction-specific improvement in motion discrimination. Vision Research 27 (6), 953–965.

Bao, S., Chang, E.F., Woods, J., Merzenich, M.M., 2004. Temporal plasticity in the primary auditory cortex induced by operant perceptual learning. Nature Neuroscience 7 (9), 974–981, http://dx.doi.org/10.1038/nn1293.

Bartolo, R., Merchant, H., 2009. Learning and generalization of time production in humans: rules of transfer across modalities and interval durations. Experimental Brain Research 197 (1), 91–100, http://dx.doi.org/10.1007/s00221-009-1895-1.

Batra, R., Kuwada, S., Fitzpatrick, D.C., 1997. Sensitivity to interaural temporal disparities of low- and high-frequency neurons in the superior olivary complex. I. Heterogeneity of responses. Journal of Neurophysiology 78 (3), 1222–1236.

Beard, B.L., Levi, D.M., Reich, L.N., 1995. Perceptual learning in parafoveal vision. Vision Research 35 (12), 1679–1690.

Beauchamp, M.S., 2005. See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. Current Opinion in Neurobiology 15 (2), 145–153, http://dx.doi.org/10.1016/j.conb.2005.03.011.

Beauchamp, M.S., Nath, A.R., Pasalar, S., 2010. fMRI-guided transcranial magnetic stimulation reveals that the superior temporal sulcus is a cortical locus of the McGurk effect. Journal of Neuroscience 30 (7), 2414–2417, http://dx.doi.org/10.1523/JNEUROSCI.4865-09.2010.

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. Cerebral Cortex 10 (5), 529–534.

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.

Bruns, P., Liebnau, R., Roder, B., 2011. Cross-modal training induces changes in spatial representations early in the auditory processing pathway. Psychological Science 22 (9), 1120–1126, http://dx.doi.org/10.1177/0956797611416254.

Calvert, G.A., Thesen, T., 2004. Multisensory integration: methodological approaches and emerging principles in the human brain. Journal of Physiology, Paris 98 (1–3), 191–205, http://dx.doi.org/10.1016/j.jphysparis.2004.03.018.

Cappe, C., Rouiller, E.M., Barone, P., 2009. Multisensory anatomical pathways. Hearing Research 258 (1–2), 28–36, http://dx.doi.org/10.1016/j.heares.2009.04.017.

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, http://dx.doi.org/10.1097/WNR.0b013e3283462def.

Cherry, E.C., 1953. Some experiments on the recognition of speech, with one and with two ears. Journal of the Acoustical Society of America 25 (5), 975–979.

Cohen, L.G., Celnik, P., Pascual-Leone, A., Corwell, B., Falz, L., Dambrosia, J., Honda, M., Sadato, N., Gerloff, C., Catala, M.D., Hallett, M., 1997. Functional relevance of cross-modal plasticity in blind humans. Nature 389 (6647), 180–183, http://dx.doi.org/10.1038/38278.

Collignon, O., Champoux, F., Voss, P., Lepore, F., 2011. Sensory rehabilitation in the plastic brain. Progress in Brain Research 191, 211–231, http://dx.doi.org/10.1016/B978-0-444-53752-2.00003-5.

De Volder, A.G., Catalan-Ahumada, M., Robert, A., Bol, A., Labar, D., Coppens, A., Michel, C., Veraart, C., 1999. Changes in occipital cortex activity in early blind humans using a sensory substitution device. Brain Research 826 (1), 128–134.

Delhommeau, K., Micheyl, C., Jouvent, R., Collet, L., 2002. Transfer of learning across durations and ears in auditory frequency discrimination. Perception and Psychophysics 64 (3), 426–436.

Dellorso, S., Schirillo, J., 2010. Light location influences the perceived locations of internal sounds. Perception 39 (1), 91–102.

Demany, L., Semal, C., 2002. Learning to perceive pitch differences. Journal of the Acoustical Society of America 111 (3), 1377–1388.

Deneve, S., Pouget, A., 2004. Bayesian multisensory integration and cross-modal spatial links. Journal of Physiology, Paris 98 (1–3), 249–258, http://dx.doi.org/10.1016/j.jphysparis.2004.03.011.

Dosher, B.A., Lu, Z.L., 1998. Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. Proceedings of the National Academy of Sciences of the United States of America 95 (23), 13988–13993.

Driver, J., Noesselt, T., 2008. Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. Neuron 57 (1), 11–23, http://dx.doi.org/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, http://dx.doi.org/10.1038/33402.

Facchini, S., Aglioti, S.M., 2003. Short term light deprivation increases tactile spatial acuity in humans. Neurology 60 (12), 1998–1999.

Fahle, M., Edelman, S., Poggio, T., 1995. Fast perceptual learning in hyperacuity. Vision Research 35 (21), 3003–3013.

Falchier, A., Clavagnier, S., Barone, P., Kennedy, H., 2002. Anatomical evidence of multimodal integration in primate striate cortex. Journal of Neuroscience 22 (13), 5749–5759.

Fenn, K.M., Nusbaum, H.C., Margoliash, D., 2003. Consolidation during sleep of perceptual learning of spoken language. Nature 425 (6958), 614–616, http://dx.doi.org/10.1038/nature01951.

Fiorentini, A., Berardi, N., 1980. Perceptual learning specific for orientation and spatial frequency. Nature 287 (5777), 43–44.

Fiorentini, A., Berardi, N., 1981. Learning in grating waveform discrimination: specificity for orientation and spatial frequency. Vision Research 21 (7), 1149–1158.

Fitzgerald, M.B., Wright, B.A., 2005. A perceptual learning investigation of the pitch elicited by amplitude-modulated noise. Journal of the Acoustical Society of America 118 (6), 3794–3803.

Frasnelli, J., Collignon, O., Voss, P., Lepore, F., 2011. Crossmodal plasticity in sensory loss. Progress in Brain Research 191, 233–249, http://dx.doi.org/10.1016/B978-0-444-53752-2.00002-3.

Ghazanfar, A.A., Schroeder, C.E., 2006. Is neocortex essentially multisensory? Trends in Cognitive Sciences 10 (6), 278–285.

Gibson, E.J., 1963. Perceptual learning. Annual Review of Psychology 14, 29–56, http://dx.doi.org/10.1146/annurev.ps.14.020163.000333.

Gilbert, C.D., Sigman, M., Crist, R.E., 2001. The neural basis of perceptual learning. Neuron 31 (5), 681–697.

Gockel, H., Colonius, H., 1997. Auditory profile analysis: is there perceptual constancy for spectral shape for stimuli roved in frequency? Journal of the Acoustical Society of America 102 (4), 2311–2315.

Goldstone, R.L., 1998. Perceptual learning. Annual Review of Psychology 49, 585–612, http://dx.doi.org/10.1146/annurev.psych.49.1.585.

Grahn, J.A., Henry, M.J., McAuley, J.D., 2011. FMRI investigation of cross-modal interactions in beat perception: audition primes vision, but not vice versa. Neuroimage 54 (2), 1231–1243, http://dx.doi.org/10.1016/j.neuroimage.2010.09.033.

Green, D.M., 1983. Profile analysis. A different view of auditory intensity discrimination. American Psychologist 38 (2), 133–142.

Grimault, N., Micheyl, C., Carlyon, R.P., Bacon, S.P., Collet, L., 2003. Learning in discrimination of frequency or modulation rate: generalization to fundamental frequency discrimination. Hearing Research 184 (1–2), 41–50.

Hervais-Adelman, A.G., Davis, M.H., Johnsrude, I.S., Taylor, K.J., Carlyon, R.P., 2011. Generalization of perceptual learning of vocoded speech. Journal of Experimental Psychology: Human Perception and Performance 37 (1), 283–295, http://dx.doi.org/10.1037/a0020772.

Holmes, N.P., Spence, C., 2005. Multisensory integration: space, time and superadditivity. Current Biology 15 (18), R762–R764, http://dx.doi.org/10.1016/j.cub.2005.08.058.

Hubel, D.H., Wiesel, T.N., 1970. The period of susceptibility to the physiological effects of unilateral eye closure in kittens. Journal of Physiology 206 (2), 419–436.

Irvine, D.R., Martin, R.L., Klimkeit, E., Smith, R., 2000. Specificity of perceptual learning in a frequency discrimination task. Journal of the Acoustical Society of America 108 (6), 2964–2968.

Iwai, E., Aihara, T., Hikosaka, K., 1987. Inferotemporal neurons of the monkey responsive to auditory signal. Brain Research 410 (1), 121–124.

Jain, A., Fuller, S., Backus, B.T., 2010. Absence of cue-recruitment for extrinsic signals: sounds, spots, and swirling dots fail to influence perceived 3D rotation direction after training. PLoS ONE 5 (10), e13295, http://dx.doi.org/10.1371/journal.pone.0013295.

Jeffress, L.A., 1948. A place theory of sound localization. Journal of Comparative and Physiological Psychology 41 (1), 35–39.

Juan, C.H., Campana, G., Walsh, V., 2004. Cortical interactions in vision and awareness: hierarchies in reverse. Progress in Brain Research 144, 117–130, http://dx.doi.org/10.1016/S0079-6123(03)14408-1.

Julesz, B., 1972. Cyclopean perception and neurophysiology. Investigative Ophthalmology 11 (6), 540–548.

Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. Journal of Neuroscience 17 (11), 4302–4311.

Karim, A.A., Schuler, A., Hegner, Y.L., Friedel, E., Godde, B., 2006. Facilitating effect of 15-Hz repetitive transcranial magnetic stimulation on tactile perceptual learning. Journal of Cognitive Neuroscience 18 (9), 1577–1585, http://dx.doi.org/10.1162/jocn.2006.18.9.1577.

Karmarkar, U.R., Buonomano, D.V., 2003. Temporal specificity of perceptual learning in an auditory discrimination task. Learning and Memory 10 (2), 141–147, http://dx.doi.org/10.1101/lm.55503.

Karni, A., Sagi, D., 1991. Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. Proceedings of the National Academy of Sciences of the United States of America 88 (11), 4966–4970.

Kauffman, T., Theoret, H., Pascual-Leone, A., 2002. Braille character discrimination in blindfolded human subjects. Neuroreport 13 (5), 571–574.

Kim, J., Kroos, C., Davis, C., 2010. Hearing a point-light talker: an auditory influence on a visual motion detection task. Perception 39 (3), 407–416.

Kim, J.K., Zatorre, R.J., 2008. Generalized learning of visual-to-auditory substitution in sighted individuals. Brain Research 1242, 263–275, http://dx.doi.org/10.1016/j.brainres.2008.06.038.

Kim, J.K., Zatorre, R.J., 2010. Can you hear shapes you touch? Experimental Brain Research 202 (4), 747–754, http://dx.doi.org/10.1007/s00221-010-2178-6.

Kim, J.K., Zatorre, R.J., 2011. Tactile–auditory shape learning engages the lateral occipital complex. Journal of Neuroscience 31 (21), 7848–7856, http://dx.doi.org/10.1523/JNEUROSCI.3399-10.2011.

King, A.J., 2009. Visual influences on auditory spatial learning. Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences 364 (1515), 331–339, http://dx.doi.org/10.1098/rstb.2008.0230.

Kraljic, T., Samuel, A.G., 2006. Generalization in perceptual learning for speech. Psychonomic Bulletin & Review 13 (2), 262–268.

Kupers, R., Fumal, A., de Noordhout, A.M., Gjedde, A., Schoenen, J., Ptito, M., 2006. Transcranial magnetic stimulation of the visual cortex induces somatotopically organized qualia in blind subjects. Proceedings of the National Academy of Sciences of the United States of America 103 (35), 13256–13260, http://dx.doi.org/10.1073/pnas.0602925103.

Kupers, R., Pietrini, P., Ricciardi, E., Ptito, M., 2011. The nature of consciousness in the visually deprived brain. Frontiers in Psychology 2, 19, http://dx.doi.org/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. Progress in Brain Research 192, 17–31, http://dx.doi.org/10.1016/B978-0-444-53355-5.00002-6.

Lapid, E., Ulrich, R., Rammsayer, T., 2009. Perceptual learning in auditory temporal discrimination: no evidence for a cross-modal transfer to the visual modality. Psychonomic Bulletin & Review 16 (2), 382–389, http://dx.doi.org/10.3758/PBR.16.2.382.

Lewald, J., 2007. More accurate sound localization induced by short-term light deprivation. Neuropsychologia 45 (6), 1215–1222, http://dx.doi.org/10.1016/j.neuropsychologia.2006.10.006.

Lewald, J., Ehrenstein, W.H., Guski, R., 2001. Spatio-temporal constraints for auditory–visual integration. Behavioural Brain Research 121 (1–2), 69–79.

Macaluso, E., Driver, J., 2005. Multisensory spatial interactions: a window onto functional integration in the human brain. Trends in Neurosciences 28 (5), 264–271, http://dx.doi.org/10.1016/j.tins.2005.03.008.

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 Research Bulletin 82 (5–6), 264–270, http://dx.doi.org/10.1016/j.brainresbull.2010.05.001.

Maye, J., Weiss, D.J., Aslin, R.N., 2008. Statistical phonetic learning in infants: facilitation and feature generalization. Developmental Science 11 (1), 122–134, http://dx.doi.org/10.1111/j.1467-7687.2007.00653.x.

McGurk, H., MacDonald, J., 1976. Hearing lips and seeing voices. Nature 264 (5588), 746–748.

McKee, S.P., Westheimer, G., 1978. Improvement in vernier acuity with practice. Perception and Psychophysics 24 (3), 258–262.

Meijer, P.B., 1992. An experimental system for auditory image representations. IEEE Transactions on Biomedical Engineering 39 (2), 112–121, http://dx.doi.org/10.1109/10.121642.

Merabet, L.B., Battelli, L., Obretenova, S., Maguire, S., Meijer, P., Pascual-Leone, A., 2009. Functional recruitment of visual cortex for sound encoded object identification in the blind. Neuroreport 20 (2), 132–138, http://dx.doi.org/10.1097/WNR.1090b1013e32832104dc.

Merabet, L.B., Hamilton, R., Schlaug, G., Swisher, J.D., Kiriakopoulos, E.T., Pitskel, N.B., Kauffman, T., Pascual-Leone, A., 2008. Rapid and reversible recruitment of early visual cortex for touch. PLoS ONE 3 (8), e3046, http://dx.doi.org/10.1371/journal.pone.0003046.

Merabet, L.B., Rizzo, J.F., Amedi, A., Somers, D.C., Pascual-Leone, A., 2005. What blindness can tell us about seeing again: merging neuroplasticity and neuroprostheses. Nature Reviews: Neuroscience 6 (1), 71–77, http://dx.doi.org/10.1038/nrn1586.

Meredith, M.A., Kryklywy, J., McMillan, A.J., Malhotra, S., Lum-Tai, R., Lomber, S.G., 2011. Crossmodal reorganization in the early deaf switches sensory, but not behavioral roles of auditory cortex. Proceedings of the National Academy of Sciences of the United States of America 108 (21), 8856–8861, http://dx.doi.org/10.1073/pnas.1018519108.

Meredith, M.A., Stein, B.E., 1986. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. Journal of Neurophysiology 56 (3), 640–662.

Micheyl, C., Bernstein, J.G., Oxenham, A.J., 2006a. Detection and F0 discrimination of harmonic complex tones in the presence of competing tones or noise. Journal of the Acoustical Society of America 120 (3), 1493–1505.

Micheyl, C., Delhommeau, K., Perrot, X., Oxenham, A.J., 2006b. Influence of musical and psychoacoustical training on pitch discrimination. Hearing Research 219 (1–2), 36–47, http://dx.doi.org/10.1016/j.heares.2006.05.004.

Murata, K., Cramer, H., Bach-y-Rita, P., 1965. Neuronal convergence of noxious, acoustic, and visual stimuli in the visual cortex of the cat. Journal of Neurophysiology 28 (6), 1223–1239.

Nagarajan, S.S., Blake, D.T., Wright, B.A., Byl, N., Merzenich, M.M., 1998. Practice-related improvements in somatosensory interval discrimination are temporally

specific but generalize across skin location, hemisphere, and modality. Journal of Neuroscience 18 (4), 1559–1570.

Noppeney, U., 2007. The effects of visual deprivation on functional and structural organization of the human brain. Neuroscience and Biobehavioral Reviews 31 (8), 1169–1180, http://dx.doi.org/10.1016/j.neubiorev.2007.04.012.

Olson, I.R., Gatenby, J.C., Gore, J.C., 2002. A comparison of bound and unbound audio-visual information processing in the human cerebral cortex. Brain Research: Cognitive Brain Research 14 (1), 129–138.

Pascual-Leone, A., Hamilton, R., 2001. The metamodal organization of the brain. Progress in Brain Research 134, 427–445.

Pasqualotto, A., Proulx, M.J., 2012. The role of visual experience for the neural basis of spatial cognition. Neuroscience and Biobehavioral Reviews 36 (4), 1179–1187, http://dx.doi.org/10.1016/j.neubiorev.2012.01.008.

Pavlovskaya, M., Hochstein, S., 2011. Perceptual learning transfer between hemispheres and tasks for easy and hard feature search conditions. Journal of Vision 11 (1), http://dx.doi.org/10.1167/11.1.8.

Petrov, A.A., Dosher, B.A., Lu, Z.L., 2005. The dynamics of perceptual learning: an incremental reweighting model. Psychological Review 112 (4), 715–743, http://dx.doi.org/10.1037/0033-295X.112.4.715.

Pietrini, P., Furey, M.L., Ricciardi, E., Gobbini, M.I., Wu, W.H., Cohen, L., Guazzelli, M., Haxby, J.V., 2004. Beyond sensory images: object-based representation in the human ventral pathway. Proceedings of the National Academy of Sciences of the United States of America 101 (15), 5658–5663, http://dx.doi.org/10.1073/pnas.0400707101.

Planetta, P.J., Servos, P., 2008. Somatosensory temporal discrimination learning generalizes to motor interval production. Brain Research 1233, 51–57, http://dx.doi.org/10.1016/j.brainres.2008.07.081.

Poggio, G.E., 1995. Mechanisms of stereopsis in monkey visual cortex. Cerebral Cortex 5 (3), 193–204.

Poirier, C., Collignon, O., Scheiber, C., Renier, L., Vanlierde, A., Tranduy, D., Veraart, C., De Volder, A.G., 2006. Auditory motion perception activates visual motion areas in early blind subjects. Neuroimage 31 (1), 279–285, http://dx.doi.org/10.1016/j.neuroimage.2005.11.036.

Pollok, B., Schnitzler, I., Stoerig, P., Mierdorf, T., Schnitzler, A., 2005. Image-to-sound conversion: experience-induced plasticity in auditory cortex of blindfolded adults. Experimental Brain Research 167 (2), 287–291, http://dx.doi.org/10.1007/s00221-005-0060-8.

Powers, A.R., Hevey, M.A., Wallace, M.T., 2012. Neural correlates of multisensory perceptual learning. The Journal of Neuroscience 32 (18), 6263–6274, http://dx.doi.org/10.1523/jneurosci.6138-11.2012.

Proulx, M.J., 2010. Synthetic synaesthesia and sensory substitution. Consciousness and Cognition 19 (1), 501–503, http://dx.doi.org/10.1016/j.concog.2009.12.005.

Proulx, M.J., Stoerig, P., 2006. Seeing sounds and tingling tongues: qualia in synaesthesia and sensory substitution. Anthropology & Philosophy 7, 135–151.

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, http://dx.doi.org/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, http://dx.doi.org/10.1093/brain/awh380.

Ptito, M., Schneider, F.C., Paulson, O.B., Kupers, R., 2008. Alterations of the visual pathways in congenital blindness. Experimental Brain Research 187 (1), 41–49, http://dx.doi.org/10.1007/s00221-008-1273-4.

Ragert, P., Schmidt, A., Altenmuller, E., Dinse, H.R., 2004. Superior tactile performance and learning in professional pianists: evidence for meta-plasticity in musicians. European Journal of Neuroscience 19 (12), 473–478.

Renier, L., Collignon, O., Poirier, C., Tranduy, D., Vanlierde, A., Bol, A., Veraart, C., De Volder, A.G., 2005. Cross-modal activation of visual cortex during depth perception using auditory substitution of vision. Neuroimage 26 (2), 573–580, http://dx.doi.org/10.1016/j.neuroimage.2005.01.047.

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, http://dx.doi.org/10.1016/j.neuron.2010.09.021.

Ricciardi, E., Pietrini, P., 2011. New light from the dark: what blindness can teach us about brain function. Current Opinion in Neurology 24 (4), 357–363, http://dx.doi.org/10.1097/WCO.0b013e328348bdbf.

Rockland, K.S., Ojima, H., 2003. Multisensory convergence in calcarine visual areas in macaque monkey. International Journal of Psychophysiology 50 (1–2), 19–26.

Ross, L.A., Saint-Amour, D., Leavitt, V.M., Javitt, D.C., Foxe, J.J., 2007. Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. Cerebral Cortex 17 (5), 1147–1153, http://dx.doi.org/10.1093/cercor/bhl024.

Saarinen, J., Levi, D.M., 1995. Perceptual learning in vernier acuity: what is learned? Vision Research 35 (4), 519–527.

Sathian, K., 2005. Visual cortical activity during tactile perception in the sighted and the visually deprived. Developmental Psychobiology 46 (3), 279–286, http://dx.doi.org/10.1002/dev.20056.

Sathian, K., Zangaladze, A., 1998. Perceptual learning in tactile hyperacuity: complete intermanual transfer but limited retention. Experimental Brain Research 118 (1), 131–134.

Sathian, K., Zangaladze, A., 2002. Feeling with the mind's eye: contribution of visual cortex to tactile perception. Behavioural Brain Research 135 (1–2), 127–132.

Schoups, A.A., Vogels, R., Orban, G.A., 1995. Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. Journal of Physiology 483 (Pt 3), 797–810.

Scott, S.K., Johnsrude, I.S., 2003. The neuroanatomical and functional organization of speech perception. Trends in Neurosciences 26 (2), 100–107.

Seitz, A.R., Kim, R., Shams, L., 2006. Sound facilitates visual learning. Current Biology 16 (14), 1422–1427, http://dx.doi.org/10.1016/j.cub.2006.05.048.

Seitz, A.R., Watanabe, T., 2009. The phenomenon of task-irrelevant perceptual learning. Vision Research 49 (21), 2604–2610, http://dx.doi.org/10.1016/j.visres.2009.08.003.

Shams, L., Seitz, A.R., 2008. Benefits of multisensory learning. Trends in Cognitive Sciences 12 (11), 411–417, http://dx.doi.org/10.1016/j.tics.2008.07.006.

Simner, J., Haywood, S.L., 2009. Tasty non-words and neighbours: the cognitive roots of lexical-gustatory synaesthesia. Cognition 110 (2), 171–181, http://dx.doi.org/10.1016/j.cognition.2008.11.008.

Small, D.A., 2004. Crossmodal integration—insights from the chemical senses. Trends in Neurosciences 27 (3), 120–123, discussion 123–124.

Stein, B.E., Burr, D., Constantinidis, C., Laurienti, P.J., Alex Meredith, M., Perrault Jr., T.J., Ramachandran, R., Roder, B., Rowland, B.A., Sathian, K., Schroeder, C.E., Shams, L., Stanford, T.R., Wallace, M.T., Yu, L., Lewkowicz, D.J., 2010. Semantic confusion regarding the development of multisensory integration: a practical solution. European Journal of Neuroscience 31 (10), 1713–1720, http://dx.doi.org/10.1111/j.1460-9568.2010.07206.x.

Sumby, W.H., Pollack, I., 1954. Visual contribution to speech intelligibility in noise. Journal of the Acoustical Society of America 26 (2), 212–215.

Tompa, T., Sary, G., 2010. A review on the inferior temporal cortex of the macaque. Brain Research Reviews 62 (2), 165–182, http://dx.doi.org/10.1016/j.brainresrev.2009.10.001.

Tsushima, Y., Seitz, A.R., Watanabe, T., 2008. Task-irrelevant learning occurs only when the irrelevant feature is weak. Current Biology 18 (12), R516–R517, http://dx.doi.org/10.1016/j.cub.2008.04.029.

van der Linden, M., van Turennout, M., Fernandez, G., 2011. Category training induces cross-modal object representations in the adult human brain. Journal of Cognitive Neuroscience 23 (6), 1315–1331, http://dx.doi.org/10.1162/jocn.2010.21522.

Vogels, R., Orban, G.A., 1985. The effect of practice on the oblique effect in line orientation judgments. Vision Research 25 (11), 1679–1687.

von Kriegstein, K., Smith, D.R., Patterson, R.D., Ives, D.T., Griffiths, T.D., 2007. Neural representation of auditory size in the human voice and in sounds from other resonant sources. Current Biology 17 (13), 1123–1128, http://dx.doi.org/10.1016/j.cub.2007.05.061.

Wang, R., Zhang, J.Y., Klein, S.A., Levi, D.M., Yu, C., 2012. Task relevancy and demand modulate double-training enabled transfer of perceptual learning. Vision Research 61, 33–38, http://dx.doi.org/10.1016/j.visres.2011.07.019.

Ward, J., Meijer, P., 2010. Visual experiences in the blind induced by an auditory sensory substitution device. Consciousness and Cognition 19 (1), 492–500, http://dx.doi.org/10.1016/j.concog.2009.10.006.

Wong, M., Gnanakumaran, V., Goldreich, D., 2011. Tactile spatial acuity enhancement in blindness: evidence for experience-dependent mechanisms. Journal of Neuroscience 31 (19), 7028–7037, http://dx.doi.org/10.1523/JNEUROSCI.6461-10.2011.

Wong, P.C., Skoe, E., Russo, N.M., Dees, T., Kraus, N., 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nature Neuroscience 10 (4), 420–422, http://dx.doi.org/10.1038/nn1872.

Wright, B.A., Buonomano, D.V., Mahncke, H.W., Merzenich, M.M., 1997. Learning and generalization of auditory temporal-interval discrimination in humans. Journal of Neuroscience 17 (10), 3956–3963.

Wright, B.A., Fitzgerald, M.B., 2001. Different patterns of human discrimination learning for two interaural cues to sound-source location. Proceedings of the National Academy of Sciences of the United States of America 98 (21), 12307–12312, http://dx.doi.org/10.1073/pnas.211220498.

Wright, B.A., Wilson, R.M., Sabin, A.T., 2010. Generalization lags behind learning on an auditory perceptual task. Journal of Neuroscience 30 (35), 11635–11639, http://dx.doi.org/10.1523/JNEUROSCI.1441-10.2010.

Wright, B.A., Zhang, Y., 2009. A review of the generalization of auditory learning. Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences 364 (1515), 301–311, http://dx.doi.org/10.1098/rstb.2008.0262.

Xiao, L.Q., Zhang, J.Y., Wang, R., Klein, S.A., Levi, D.M., Yu, C., 2008. Complete transfer of perceptual learning across retinal locations enabled by double training. Current Biology 18 (24), 1922–1926, http://dx.doi.org/10.1016/j.cub.2008.10.030.

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, http://dx.doi.org/10.1038/44139.

Zatorre, R.J., Bouffard, M., Belin, P., 2004. Sensitivity to auditory object features in human temporal neocortex. Journal of Neuroscience 24 (14), 3637–3642, http://dx.doi.org/10.1523/JNEUROSCI.5458-03.2004.

Zatorre, R.J., Evans, A.C., Meyer, E., Gjedde, A., 1992. Lateralization of phonetic and pitch discrimination in speech processing. Science 256 (5058), 846–849.

Zhang, J.Y., Kuai, S.G., Xiao, L.Q., Klein, S.A., Levi, D.M., Yu, C., 2008. Stimulus coding rules for perceptual learning. PLoS Biology 6 (8), e197, http://dx.doi.org/10.1371/journal.pbio.0060197.

Zhang, J.Y., Zhang, G.L., Xiao, L.Q., Klein, S.A., Levi, D.M., Yu, C., 2010. Rule-based learning explains visual perceptual learning and its specificity and transfer. Journal of Neuroscience 30 (37), 12323–12328, http://dx.doi.org/10.1523/JNEUROSCI.0704-10.2010.