

Benefits of multisensory learning

Ladan Shams¹ and Aaron R. Seitz²

¹ Department of Psychology, University of California, Los Angeles, CA 90095, USA

² Department of Psychology, University of California, Riverside, CA 92521, USA

Studies of learning, and in particular perceptual learning, have focused on learning of stimuli consisting of a single sensory modality. However, our experience in the world involves constant multisensory stimulation. For instance, visual and auditory information are integrated in performing many tasks that involve localizing and tracking moving objects. Therefore, it is likely that the human brain has evolved to develop, learn and operate optimally in multisensory environments. We suggest that training protocols that employ unisensory stimulus regimes do not engage multisensory learning mechanisms and, therefore, might not be optimal for learning. However, multisensory-training protocols can better approximate natural settings and are more effective for learning.

Suppose that an adult individual wishes to learn to discriminate a variety of bird species. What would be the best training protocol for this kind of learning? Acquiring this skill can involve many types of learning and here we focus on aspects of perceptual learning; namely, improvements in the perceptual skills that enable recognition and discrimination of the birds. Based on many studies of visual perceptual learning [1–5], we can hypothesize that this learning might be obtained through training with pictures or video clips of the bird types labeled with the name of the species. However, mastering perceptual skills is often slow and effortful in adults (although some initial learning can be rapid [6]). Studies of perceptual learning show that a simple task of detecting a subtle visual stimulus can require a month or more of training to asymptote, with only a small increment of performance [7–9].

Would the learning be easier if the training combined the images with the sound of the birds? Most likely ‘yes’ because the bird songs are another feature that can be used to distinguish the different species. What if the individual must discriminate among the species only based on pictures in the absence of any sound? Would training with auditory–visual (i.e. multisensory) stimuli still be beneficial if the ultimate goal of learning is to visually discriminate them? Traditionally, researchers of perception would answer ‘no’, with the assumption that if the task to be learned is visual, then sound features that would be absent during the task would not be very helpful for learning, and could even be disruptive to learning by distracting attention away from visual features. However, here we discuss findings that show the opposite.

We note that the human brain has evolved to learn and operate in natural environments in which behavior is often guided by information integrated across multiple sensory modalities. Multisensory interactions are ubiquitous in the nervous system and occur at early stages of perceptual processing. Therefore, unisensory-training protocols used for skill acquisition in adults can provide unnatural settings and do not tap into multisensory learning mechanisms that have evolved to produce optimal behavior in the naturally multisensory environment. Within this framework, we argue that multisensory-training protocols, as opposed to unisensory protocols, can better approximate natural settings and, therefore, produce greater and more efficient learning. However, the extent to which this facilitation occurs depends upon appropriate relations (i.e. congruency; between the information coming into each of the senses).

The multisensory brain

Perception has traditionally been viewed as a modular function with the different sensory modalities operating largely as separate and independent processes. However, an overwhelming set of new findings has overturned this dogma. Reports of multisensory interactions in various perceptual tasks and settings indicate that these interactions are the rule rather than the exception in human processing of sensory information [10,11] and there exists a rapidly growing literature of the neuroanatomical, electrophysiological and neuroimaging studies that show that multisensory interactions can occur throughout processing [11–13].

There are numerous brain areas and pathways for multisensory interactions, ranging from the brain stem [14] to early sensory cortical areas [13,15], to association and other cortical areas [12], including feedforward [16] and feedback pathways [17,18]. Of particular interest are recently discovered multisensory modulations, activations and connectivity at the earliest stages of perceptual processing; areas that have long been viewed as ‘sensory-specific’ [12,13] (also see Refs [19–21]). Such findings of multisensory interactions in early brain areas, raise the question of whether any brain regions can be fully characterized through their unisensory response properties [12], and motivate our assertion that learning at all processing stages can involve multisensory processes.

Multisensory plasticity in development

Both animal and human studies of the early stages of life show that, during early development, an alteration to the environment or a disruption of processing to one sense, can result in a striking degree of neural plasticity between the senses [10,11,22–25]. For example, changing spatial

Corresponding authors: Shams, L. (ladan@psych.ucla.edu); Seitz, A.R. (aseitz@ucr.edu).

correlations between auditory and visual inputs alters the multisensory representation in the superior colliculus [26], and disruption to the inputs to auditory cortex can result in this area being driven by visual inputs [27]. These findings provide evidence that learning and plasticity in early life can be highly multisensory.

Furthermore, multisensory stimulation has been argued to provide a redundancy that is crucial for extracting information that would not be possible based on unisensory stimulation alone in early infancy. For example, Bahrick and Lickliter [28] showed that 5-month-old infants could discriminate visually presented rhythms only if they were habituated with auditory–visual presentations of the rhythm and not when habituated with visual-only or auditory-only presentations of rhythm.

Multisensory facilitation of unisensory learning

Although sensory plasticity [5] and perceptual learning in adults is considerably more restricted compared to development [29], accumulating reports indicate the superiority of bisensory training in adult learners. A recent study compared auditory–visual and visual training for perceptual learning by using a coherent motion detection and discrimination task [30]. Compared to a visual (V) group, the auditory–visual (AV) group showed greater learning both within the first session (Figure 1a) and across the ten training sessions (Figure 1b). The two groups were compared on trials that only contained visual signals and no auditory signal and, therefore, the advantage of multisensory training was evident even in the absence of auditory information. The advantage of auditory–visual training over visual-alone training in this study was substantial; it reduced the number of sessions required to reach asymptote by as much as ~60%, while also raising the maximum performance.

Complementary results were found in a study of auditory recognition [31]. Voice recognition was improved by audiovisual training (voice co-presented with video of the speaking face) as compared to training with voice alone.

Multisensory exposure has also been reported to enhance reinforcement learning [32] in *Drosophila* (e.g. fruit flies). *Drosophila* conditioned with concurrent visual and olfactory

cues demonstrated facilitated learning of subthreshold stimuli as compared to *drosophila* conditioned with only visual cues. More importantly, unisensory memory retrieval was also enhanced by the multisensory learning conditions. Moreover, preconditioning with bimodal stimuli followed by unisensory conditioning led to crossmodal memory transfer.

These results indicate that multisensory training promotes more effective learning of the information than unisensory training. Although these findings span a large range of processing levels and might be mediated by different mechanisms, it nonetheless seems that the multisensory benefit to learning is an overarching phenomenon.

The importance of congruency

Can training on any arbitrary pair of auditory and visual stimuli have such facilitatory learning effects? We suspect that the answer is ‘yes and no’. Although training on any pair of multisensory stimuli might induce a more effective representation of the unisensory stimulus, the effects could be substantially more pronounced for congruent stimuli. Here, we define congruency broadly as the relationship between stimuli that are consistent with the prior experience of the individual or relationships between the senses found in nature. This spans the basic attributes such as concordance in space and time, in addition to higher-level features such as semantic content (e.g. object and speech information). Although some of these relationships might be more fundamental than others, we regard congruency as an overall function of all these relationships. As such, the extant data seem to support the role of congruency in learning facilitation.

For example, Kim, Seitz and Shams [33] compared learning across three groups, one trained with visual motion stimuli only, one with congruent auditory–visual (both moving in the same direction) and one with incongruent auditory–visual stimuli (moving in opposite directions). They found that facilitation was specific to the congruent condition (Figure 1c), and concluded that facilitation involves multisensory interactions (as opposed to task-relevant information or a general alerting effect induced by sound) because the results cannot be easily explained without a consideration of the crossmodal featural relationships.

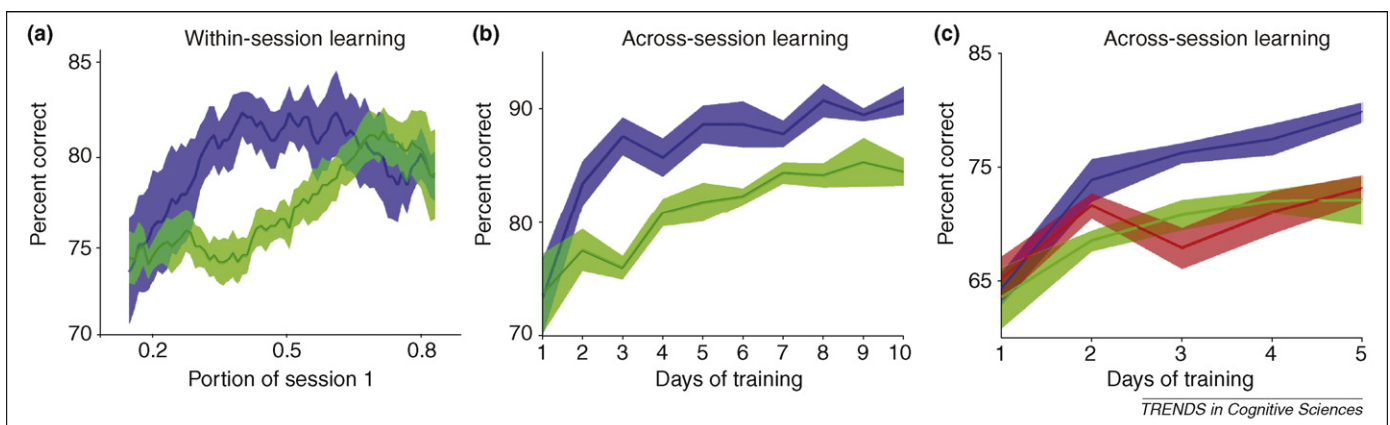


Figure 1. Results of multisensory facilitation of visual learning. Congruent (blue) auditory–visual training results in, (a) faster within-session learning and (b) greater across-session learning, compared to (c) unisensory training (red). Incongruent (green) training does not show same benefit as congruent training (blue). Data in all plots is for trials containing only visual signals. Shaded regions indicate within subject standard error. Parts (a) and (b) reproduced, with permission, from Ref. [30]. Part (c) adapted from Ref. [33].

Box 1. Acquisition of new multisensory associations

Recent studies using a diversity of tasks, species and paradigms have provided evidence that novel and arbitrary associations between multisensory features can be acquired in the mature brain. Studies have found that primates can learn multisensory associations, and in the case of haptic–visual associations, multimodal areas such as the Anterior Rhinal Cortex [41,42] are necessary for this learning. However, Gibson and Maunsell [43] found that macaque monkeys learned to associate arbitrary auditory and visual stimuli and, after this training, cells in inferotemporal cortex showed responses selective to the trained multisensory associations, indicating that multisensory learning also alters responses in brain areas often considered as unisensory. Multisensory learning is not limited to primates as shown by Guo and Guo [32] who found that conditioning in *Drosophila* was formed between arbitrary visual and olfactory cues.

Recent behavioral studies in humans show that novel multisensory associations can develop very rapidly. For example, Seitz *et al.* [5] found that after an 8 min exposure period to a rapid serial

presentation of auditory and visual stimuli, subjects learned to identify auditory–visual ‘objects’ that showed statistical regularities within the sequence. This type of statistical learning has been argued to be the basic process of new object formation in both infants and adults [44]. In a related paradigm, Kawahara [45] found that novel auditory–visual contexts could be implicitly learned by human subjects. Tanabe *et al.* [46] investigated neural substrates of audio-visual paired-association learning with functional Magnetic Resonance Imaging (fMRI) during a delayed matching to sample task. As learning progressed, visual cortical activity increased during the delay after an auditory component of a pair was presented, and conversely, superior temporal gyrus displayed more activity in response to the visual component of a pair with learning. Whereas these studies examined high-level object formation, a few recent studies have shown that association between arbitrary low-level auditory and visual [47] and visual and haptic [48] features can also be learned in human adults.

Similarly, von Kriegstein and Giraud [31] reported that although the recognition of cell phone rings was improved after arbitrary pairing with the cell phone images, this improvement was much smaller compared to that of voices paired with faces.

Consistent with these results, research has demonstrated that low-level multisensory interactions are specific to congruent stimulus features (e.g. Ref. [34]). Although congruency is clearly important for learning, it is worth

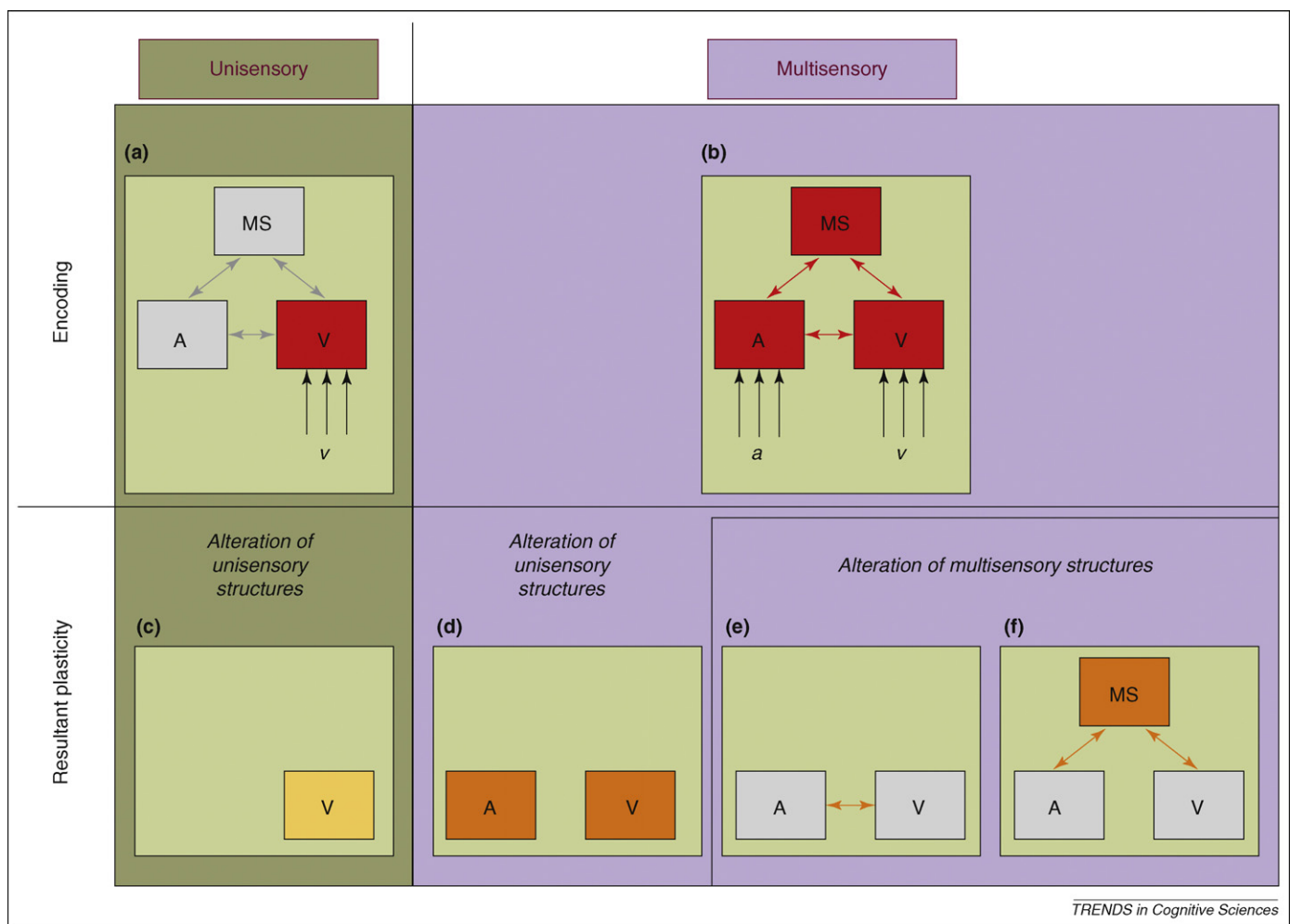


Figure 2. Schematics of processes of multisensory versus unisensory learning. (a) During encoding for unisensory training, only visual inputs (*v*) are present and only visual structures are activated (shown in red). (b) In multisensory training, both visual (*v*) and auditory (*a*) inputs are present and a larger set of processing structures are activated. A, V, and MS represent auditory, visual, and multisensory areas, respectively. (c),(d),(e) and (f) Different alterations that can result from learning. Areas and connections that undergo learning are shown in orange, light orange and bright orange represent low and high degrees of plasticity, respectively. (c) With unisensory training, some alteration can be achieved in corresponding unisensory structures (d) or in multisensory structures (e) and (f).

noting that many multisensory congruencies can only exist as a function of our experience. For example, congruencies such as a dog making a ‘woof’ sound and a bell making ‘dong’ sound are experientially driven. Given this, it is possible that if training continued for a longer period of time, the facilitation by the arbitrary cell phone and ring type in von Kriegstein and Giraud [31] might have become equally effective in the subsequent recognition task. Indeed, new evidence indicates that arbitrary associations between cross modal features can be learned in adult observers (Box 1), although less clear is the extent to which learning can overcome relationships that are prevalent in nature such as the incongruent motion directions used by Kim, Seitz and Shams [30].

Mechanisms of multisensory facilitation of unisensory learning

How does multisensory training benefit later performance of unisensory tasks? The answer must involve an alteration in the representation of the unisensory information. The key difference between unisensory and multisensory training exists during encoding, whereby a larger set of processing structures are activated in the multisensory paradigms (Figure 2a,b). However, there are several possibilities regarding which structures are altered during this learning process. A dichotomy can be found between frameworks that conjecture that multisensory processing facilitates changes within unisensory representations (Figure 2d), and those that conjecture alteration or formation of multisensory representations (Figure 2e,f). In other words, in one framework the learning affects the same structures and representations that undergo modification using the classic unisensory training (Figure 2d) and the result of facilitation is that later presentation of unisensory stimuli will yield a stronger activation of unisensory structures. Whereas in the other, the learning involves alteration of connections between modalities (Figure 2e) or the formation or alteration of multisensory representations (Figure 2f), and the later presentation of unisensory stimuli will activate a wider, multisensory, network of brain regions.

How can the learning only affect ‘unisensory’ (e.g. visual) structures? One possibility is that activation of

neurons of one modality is modulated by the signals originating from neurons of a different modality. It has been proposed that for low-level perceptual learning to occur, a threshold in neural activation needs to be exceeded [5]. This theory would predict that by modulating the activity in visual areas, sound could cause the visual activation to exceed the threshold and, thus, facilitate visual learning (or vice versa). Some studies of perceptual learning have found specificity of learning to primitive sensory features (such as spatial location, angle of orientation or motion directions), indicating that learning can take place in early stages of sensory processing [2]. The finding of auditory facilitation of visual motion learning (discussed earlier) fits this framework, in that training resulted in sensitivity changes in simple motion-detection tasks and these were found to be specific to trained visual motion direction [30,33]. Thus, Seitz and Dinse [5] accounted for this phenomenon as an example of how multisensory stimulation could boost visual (auditory) processing and result in alterations of unisensory visual (and/or auditory) structures (Figure 2d). Multisensory enhancement of activity can take on several forms, including increases in the firing rate [14,35], resetting the phase of ongoing oscillatory activity [36], decreasing response latencies [37,38] and each of these mechanisms could have the effect of enhancing plasticity.

Alternatively, facilitation can be explained by multisensory exposure resulting in formation or alteration of a multisensory representation that can then be invoked by a unisensory component (Figure 2e,f). This conjecture has been largely motivated by Predictive Coding models [39,40]. The auditory facilitation of visual motion learning discussed earlier is also consistent with this framework. Although visual motion activates primarily the visual motion processing areas, repeated auditory–visual motion presentation could result in alterations in the functional connectivity between unisensory areas (Figure 2e), or in alterations within (or connectivity with) a multisensory area that receives input from relevant unisensory regions (Figure 2f). As a result, even the visual motion alone would activate auditory or multisensory motion processing areas, hence providing a richer representation upon which detection and discrimination processes can operate. Some recent

Box 2. Multisensory facilitation of memory

Memory research shows that multisensory exposure can result in superior recognition of objects compared to unisensory exposure. For example, it has been suggested that auditory–visual synaesthesia can provide a superior memory capacity [49]. Likewise, Lehman and Murray [50] had subjects conduct a recognition task in which subjects indicated for each stimulus presentation whether that stimulus had been previously presented; a subset of these stimuli were paired with sounds that corresponded to the image identity (e.g. image of a bell and ‘dong’). Even though sound was not present in the second presentation, they found that the images that were previously presented with their corresponding sound were recognized better than those that had previously been presented only visually. Similar to the case of perceptual learning, congruency (here semantic congruency) has an important role in multisensory facilitation and incongruent stimuli (e.g. image of a bell and a ‘woof’ sound) did not facilitate memory.

Results of this kind are in line with the classic psychology notion of ‘reintegration’, which refers to the phenomenon in which a whole

state of mind is restored from an element of the whole [51–54]. Several recent neuroimaging studies of memory have provided results that shed light on the mechanism of reintegration. An fMRI study of object memory showed that recognition of images activated olfactory (piriform) cortex for images that had initially been presented in the context of an odor [55]. A Positron emission tomography study of word recognition found that the visual recognition of words that were encoded using auditory and visual representations activates auditory cortical areas, even though the recognition did not require retrieval of auditory information [53]. However, the role of imagery in these findings is not entirely ruled out, and it is also not clear whether the multisensory activations indeed contribute to the retrieval process or are epiphenomenal. Nonetheless, these and other studies [56,57] suggest that multisensory exposure enables stimuli to be encoded into multisensory representations and, thus, will later activate a larger network of brain areas than those invoked after unisensory encoding, and this might underlie the behavioral facilitation observed in some studies [50].

Box 3. Multisensory educational methods

Educators and clinicians have long believed that multisensory training can enhance learning [58–61]. A simple advantage of multisensory training is that it can engage individuals with different learning styles, for example, some people are ‘visual learners’ and others ‘auditory learners’. However, above and beyond this, multisensory training is demonstratively more effective at an individual level. For example, Treichler [62] stated ‘People generally remember 10% of what they read, 20% of what they hear, 30% of what they see, and 50% of what they see and hear’. Montessori [59,63] started the multisensory learning movement approximately ninety years ago, and most subject areas in a Montessori school use a mixture of visual, auditory, tactile and kinesthetic approaches. Over the last 1.5 decades, several modern language instruction techniques have coalesced into a method called Multisensory Structural Language Education, which uses visual, auditory, tactile-kinesthetic and articulatory modalities for teaching [64].

With the advent of technology in the classroom, a growing research area in multimedia educational techniques strongly parallels percep-

tual research of multisensory facilitation. The principle of ‘dual coding’ [65] indicates that information entering the system through multiple processing channels helps circumvent the limited processing capabilities of each individual channel and, thus, greater total information can be processed when spread between multiple senses. Related research indicates that multimodal processing reduces cognitive load because information from different modalities can be more easily chunked into short-term memory and used to build long-term representations [66]. The ‘Cognitive Theory of Multimedia Learning’ [67] builds upon this idea through studies of classroom materials; animations that are co-presented with narrations facilitate learning of facts and concepts. Importantly, this framework indicates key constraints on how this facilitation of learning occurs [68]. Namely, that additional information within the same modality, which saturates one channel with information, and materials that are incoherent with the core concept interfere with learning. Altogether, these findings indicate that research on the mechanisms of multisensory facilitation can have important benefits in pedagogy.

findings have been reported to support this mechanism. For example, von Kriegstein and Giraud [31] reported that functional connectivity between face (e.g. fusiform face area) and voice (e.g. temporal voice areas) areas is increased after voice–face training (Figure 2e) and these associative representations subsequently became available for unisensory recognition. Several recent studies of memory are consistent with this framework as well (Box 2).

Further research is needed to discriminate between accounts of learning that suggest alterations of unisensory structures (Figure 2d), those suggesting alterations in connectivity between unisensory structures (Figure 2e) and those suggesting alterations of multisensory structures (Figure 2f). However, it is important to note that these forms of learning are not mutually exclusive and can occur in parallel. Given that multisensory interactions occur at multiple stages of processing, it is likely that all three types of mechanisms are operating in the brain, and the relative contribution of each depends on the task, and factors such as stimuli and extent of training.

Relevance of the findings

These findings indicate that learning mechanisms operate optimally under multisensory conditions. But does this finding have practical relevance? Are there many conditions

in which training can be made multisensory with meaningful congruent stimuli, or is this finding only applicable to very few learning situations? This is of course an empirical question, but we speculate that these findings can have important implications for real-life learning applications. Most objects have multisensory representations. Most abstract notions can be communicated using multiple media (e.g. by written words or images, by auditory articulation or their natural sounds). In fact, some of the most successful education philosophies have incorporated multisensory training into their curricula (Box 3). These educators have empirically noticed that multisensory training for various tasks is more effective than the traditional unisensory training schemes, and have incorporated these approaches into educational protocols at all levels of instruction.

Conclusions

The results presented here demonstrate that multisensory training can be more effective than similar unisensory-training paradigms. However, this facilitation is limited by preexisting congruencies of information coming from the different senses. In general, we conjecture that perceptual and cognitive mechanisms have evolved for, and are tuned to, processing multisensory signals. Under such a regime, encoding, storing and retrieving perceptual information is intended by default to operate in a multisensory environment, and unisensory processing is often suboptimal as it would correspond to an artificial mode of processing that does not use the perceptual machinery to its fullest potential. However, research of learning is typically conducted with simple stimuli in minimal settings. We encourage more research on multisensory learning paradigms so that the mechanisms and processes of learning within natural settings can be better understood (Box 4).

Acknowledgement

We thank Bonnie Hasse for bringing the multisensory training aspects of the Montessori method to our attention and Robyn Kim and Russell Poldrack for their comments on the manuscript. We are grateful to the anonymous reviewers for their insightful comments.

References

- 1 Ahissar, M. and Hochstein, S. (2004) The reverse hierarchy theory of visual perceptual learning. *Trends Cogn. Sci.* 8, 457–464

Box 4. Questions for future research

- How general is the multisensory benefit to learning? Is it restricted to certain tasks, or is it a universal property of perceptual and cognitive learning?
- What are the mechanisms by which multisensory training modulates activity during encoding, and does this involve different processes than found in unisensory training?
- To what extent does multisensory facilitation result in alterations within unisensory versus multisensory structures? How does this differ across different training procedures?
- Do long-term and/or hard-wired multisensory associations facilitate learning in a fundamentally different way than newly trained ones? For example, can adults learn a new multisensory congruency that will be equally effective as one formed during a developmental critical period?
- Are the observed advantages of multisensory training unique to stimulation from multiple modalities or do they generalize to any redundant set of sensory stimuli, even within the same modality?

- 2 Fahle, M. (2005) Perceptual learning: specificity versus generalization. *Curr. Opin. Neurobiol.* 15, 154–160
- 3 Gilbert, C.D. *et al.* (2001) The neural basis of perceptual learning. *Neuron* 31, 681–697
- 4 Seitz, A. and Watanabe, T. (2005) A unified model for perceptual learning. *Trends Cogn. Sci.* 9, 329–334
- 5 Seitz, A.R. and Dinse, H.R. (2007) A common framework for perceptual learning. *Curr. Opin. Neurobiol.* 17, 148–153
- 6 Poggio, T. *et al.* (1992) Fast perceptual learning in visual hyperacuity. *Science* 256, 1018–1021
- 7 Furmanski, C.S. *et al.* (2004) Learning strengthens the response of primary visual cortex to simple patterns. *Curr. Biol.* 14, 573–578
- 8 Schoups, A.A. *et al.* (1995) Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. *J. Physiol.* 483, 797–810
- 9 Watanabe, T. *et al.* (2002) Greater plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task. *Nat. Neurosci.* 5, 1003–1009
- 10 Shimojo, S. and Shams, L. (2001) Sensory modalities are not separate modalities: plasticity and interactions. *Curr. Opin. Neurobiol.* 11, 505–509
- 11 Calvert, G. *et al.* (2004) *The Handbook of Multisensory Processes*. MIT Press
- 12 Ghazanfar, A.A. and Schroeder, C.E. (2006) Is neocortex essentially multisensory? *Trends Cogn. Sci.* 10, 278–285
- 13 Driver, J. and Noesselt, T. (2008) Multisensory interplay reveals crossmodal influences on ‘sensory-specific’ brain regions, neural responses, and judgments. *Neuron* 57, 11–23
- 14 Stein, B.E. and Meredith, M.A. (1993) *The Merging of the Senses*. MIT Press
- 15 Ghazanfar, A. and Schroeder, C.E. (2006) Is neocortex essentially multisensory? *Trends Cogn. Sci.* 10, 278–285
- 16 Foxe, J.J. and Schroeder, C.E. (2005) The case for feedforward multisensory convergence during early cortical processing. *Neuroreport* 16, 419–423
- 17 Falchier, A. *et al.* (2002) Anatomical evidence of multimodal integration in primate striate cortex. *J. Neurosci.* 22, 5749–5759
- 18 Schroeder, C.E. and Foxe, J.J. (2002) The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res. Cogn. Brain Res.* 14, 187–198
- 19 Watkins, S. *et al.* (2007) Activity in human V1 follows multisensory perception. *Neuroimage* 37, 572–578
- 20 Watkins, S. *et al.* (2006) Sound alters activity in human V1 in association with illusory visual perception. *Neuroimage* 31, 1247–1256
- 21 Bulkin, D. *et al.* (2008) Visual information in the ascending auditory pathway. *J. Vis.* 8, 527a
- 22 Majewska, A.K. and Sur, M. (2006) Plasticity and specificity of cortical processing networks. *Trends Neurosci.* 29, 323–329
- 23 Sleigh, M.J. and Lickliter, R. (1997) Augmented prenatal auditory stimulation alters postnatal perception, arousal, and survival in bobwhite quail chicks. *Dev. Psychobiol.* 30, 201–212
- 24 Bavelier, D. and Neville, H.J. (2002) Cross-modal plasticity: where and how? *Nat. Rev. Neurosci.* 3, 443–452
- 25 Pascual-Leone, A. *et al.* (2005) The plastic human brain cortex. *Annu. Rev. Neurosci.* 28, 377–401
- 26 Wallace, M.T. and Stein, B.E. (2007) Early experience determines how the senses will interact. *J. Neurophysiol.* 97, 921–926
- 27 Sharma, J. *et al.* (2000) Induction of visual orientation modules in auditory cortex. *Nature* 404, 841–847
- 28 Bahrick, L.E. and Lickliter, R. (2000) Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Dev. Psychol.* 36, 190–201
- 29 Fahle, M. and Poggio, T. (2002) *Perceptual Learning*. MIT Press
- 30 Seitz, A.R. *et al.* (2006) Sound facilitates visual learning. *Curr. Biol.* 16, 1422–1427
- 31 von Kriegstein, K. and Giraud, A.L. (2006) Implicit multisensory associations influence voice recognition. *PLoS Biol.* 4, e326
- 32 Guo, J. and Guo, A. (2005) Crossmodal interactions between olfactory and visual learning in *Drosophila*. *Science* 309, 307–310
- 33 Kim, R.S. *et al.* (2008) Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS One* 3, e1532
- 34 Ramos-Estebanez, C. *et al.* (2007) Visual phosphene perception modulated by subthreshold crossmodal sensory stimulation. *J. Neurosci.* 27, 4178–4181
- 35 Allman, B.L. *et al.* (2008) Subthreshold auditory inputs to extrastriate visual neurons are responsive to parametric changes in stimulus quality: Sensory-specific versus non-specific coding. *Brain Res.*, DOI: 10.1016/j.brainres.2008.03.086
- 36 Lakatos, P. *et al.* (2007) Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53, 279–292
- 37 Alpert, G.F. *et al.* (2008) Temporal characteristics of audiovisual information processing. *J. Neurosci.* 28, 5344–5349
- 38 Martuzzi, R. *et al.* (2007) Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cereb. Cortex* 17, 1672–1679
- 39 Friston, K. (2005) A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 815–836
- 40 Rao, R.P. and Ballard, D.H. (1999) Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87
- 41 Goulet, S. and Murray, E.A. (2001) Neural substrates of crossmodal association memory in monkeys: the amygdala versus the anterior rhinal cortex. *Behav. Neurosci.* 115, 271–284
- 42 Murray, E.A. and Mishkin, M. (1985) Amygdectomy impairs crossmodal association in monkeys. *Science* 228, 604–606
- 43 Gibson, J.R. and Maunsell, J.H. (1997) Sensory modality specificity of neural activity related to memory in visual cortex. *J. Neurophysiol.* 78, 1263–1275
- 44 Fiser, J. and Aslin, R.N. (2001) Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychol. Sci.* 12, 499–504
- 45 Kawahara, J. (2007) Auditory-visual contextual cuing effect. *Percept. Psychophys.* 69, 1399–1408
- 46 Tanabe, H.C. *et al.* (2005) Functionally segregated neural substrates for arbitrary audiovisual paired-association learning. *J. Neurosci.* 25, 6409–6418
- 47 Michel, M.M. and Jacobs, R.A. (2007) Parameter learning but not structure learning: a Bayesian network model of constraints on early perceptual learning. *J. Vis.* 4, 1–18
- 48 Ernst, M.O. (2007) Learning to integration arbitrary signals from vision and touch. *J. Vis.* 7, 1–14
- 49 Luria, A.R. (1987) *The Mind of a Mnemonist: A Little Book About a Vast Memory*. Harvard University Press
- 50 Lehmann, S. and Murray, M.M. (2005) The role of multisensory memories in unisensory object discrimination. *Brain Res. Cogn. Brain Res.* 24, 326–334
- 51 Hamilton, W. (1859) In *Lectures on Metaphysics and Logic* (Vol. I), Gould & Lincoln
- 52 Horowitz, L.M. and Prytulak, L.S. (1969) Redintegrative memory. *Psychol. Rev.* 76, 519–532
- 53 Nyberg, L. *et al.* (2000) Reactivation of encoding-related brain activity during memory retrieval. *Proc. Natl. Acad. Sci. U. S. A.* 97, 11120–11124
- 54 Tulving, E. and Madigan, S.A. (1970) Memory and verbal learning. *Annu. Rev. Psychol.* 21, 437–484
- 55 Gottfried, J.A. *et al.* (2004) Remembrance of odors past: human olfactory cortex in cross-modal recognition memory. *Neuron* 42, 687–695
- 56 Murray, M.M. *et al.* (2004) Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *Neuroimage* 21, 125–135
- 57 Murray, M.M. *et al.* (2005) The brain uses single-trial multisensory memories to discriminate without awareness. *Neuroimage* 27, 473–478
- 58 Fernald, G.M. and Keller, H. (1921) The effect of kinesthetic factors in development of word recognition in the case of non-readers. *J. Educ. Res.* 4, 355–377
- 59 Montessori, M. (1912) *The Montessori Method*. Frederick Stokes
- 60 Orton, S.T. (1928) Specific reading disability—strephosymbolia. *J. Am. Med. Assoc.* 90, 1095–1099
- 61 Strauss, A. and Lehtinen, L.E. (1947) *Psychopathology and Education of the Brain-Injured Child*. Grune & Stratton
- 62 Treichler, D.G. (1967) Are you missing the boat in training aid? *Film AV Commun.* 1, 14–16
- 63 Montessori, M. (1967) *The Absorbant Mind*, Holt, Rinehart and Winston

- 64 Birsh, J.R. (1999) *Multisensory Teaching of Basic Language Skills*. Paul H. Brookes Publishing Co
- 65 Clark, J.M. and Paivio, A. (1991) Dual coding theory and education. *Educ. Psychol.* 37, 250–263
- 66 Bagui, S. (1998) Reasons for increased learning using multimedia. *J. Educ. Multimed. Hypermedia* 7, 3–18
- 67 Harp, S.F. and Mayer, R.E. (1998) How seductive details do their damage: a theory of cognitive interest in science learning. *J. Educ. Psychol.* 93, 187–198
- 68 Mayer, R.E. *et al.* (2001) Cognitive constraints on multimedia learning: when presenting more material results in less understanding. *J. Educ. Psychol.* 93, 187–198