



Influences of multisensory experience on subsequent unisensory processing

Ladan Shams^{1,2*}, David R. Wozny^{2,3}, Robyn Kim¹ and Aaron Seitz⁴

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

² Biomedical Engineering Interdepartmental Program, University of California Los Angeles, Los Angeles, CA, USA

³ Department of Otolaryngology, Oregon Health and Science University, Portland, OR, USA

⁴ Department of Psychology, University of California Riverside, Riverside, CA, USA

Edited by:

Nadia Bolognini, University
Milano-Bicocca, Italy

Reviewed by:

Thomas J. Perrault Jr., Wake Forest
University, USA
Francesco Marini, University of
Milano-Bicocca, Italy

***Correspondence:**

Ladan Shams, Department of
Psychology, University of California
Los Angeles, Los Angeles, CA
90095-1563, USA.
e-mail: ladan@psych.ucla.edu

Multisensory perception has been the focus of intense investigation in recent years. It is now well-established that crossmodal interactions are ubiquitous in perceptual processing and endow the system with improved precision, accuracy, processing speed, etc. While these findings have shed much light on principles and mechanisms of perception, ultimately it is not very surprising that multiple sources of information provides benefits in performance compared to a single source of information. Here, we argue that the more surprising recent findings are those showing that multisensory experience also influences the subsequent *unisensory* processing. For example, exposure to auditory–visual stimuli can change the way that auditory or visual stimuli are processed subsequently even in isolation. We review three sets of findings that represent three different types of learning ranging from perceptual learning, to sensory recalibration, to associative learning. In all these cases exposure to multisensory stimuli profoundly influences the subsequent unisensory processing. This diversity of phenomena may suggest that continuous modification of unisensory representations by multisensory relationships may be a general learning strategy employed by the brain.

Keywords: **multisensory integration, multisensory representation, unisensory representation, multisensory learning, learning facilitation**

INTRODUCTION

We live in a world that is replete with multisensory information. As such, multisensory processing has been an active topic of research and numerous studies have demonstrated that multisensory processing can improve accuracy (e.g., Sumby and Pollack, 1954, reduce reaction times, e.g., Gingras et al., 2009), improve precision (e.g., Ernst and Banks, 2002; Alais and Burr, 2004), and provide more complete information about objects (Newell et al., 2001). Furthermore, recent studies have established the presence of a significant degree of plasticity in multisensory processes, including processes such as crossmodal simultaneity (e.g., Fujisaki et al., 2004, and temporal order, e.g., Miyazaki et al., 2006) that had previously been thought to be hardwired or highly stable. However, how multisensory processing impacts subsequent *unisensory* processing has received less attention. This is despite the fact that several studies indicate that unisensory processing is altered through multisensory experience.

In Section “Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training,” we describe recent studies that show that training observers using correlated auditory–visual stimuli improves subsequent performance in a unisensory (visual or auditory) detection, discrimination, and recognition task. In Section “Change in Unisensory Map as a Result of Exposure to Crossmodal Error,” we discuss recent research demonstrating that momentary exposure to auditory–visual spatial discrepancy results in a shift in the auditory space map. We discuss how

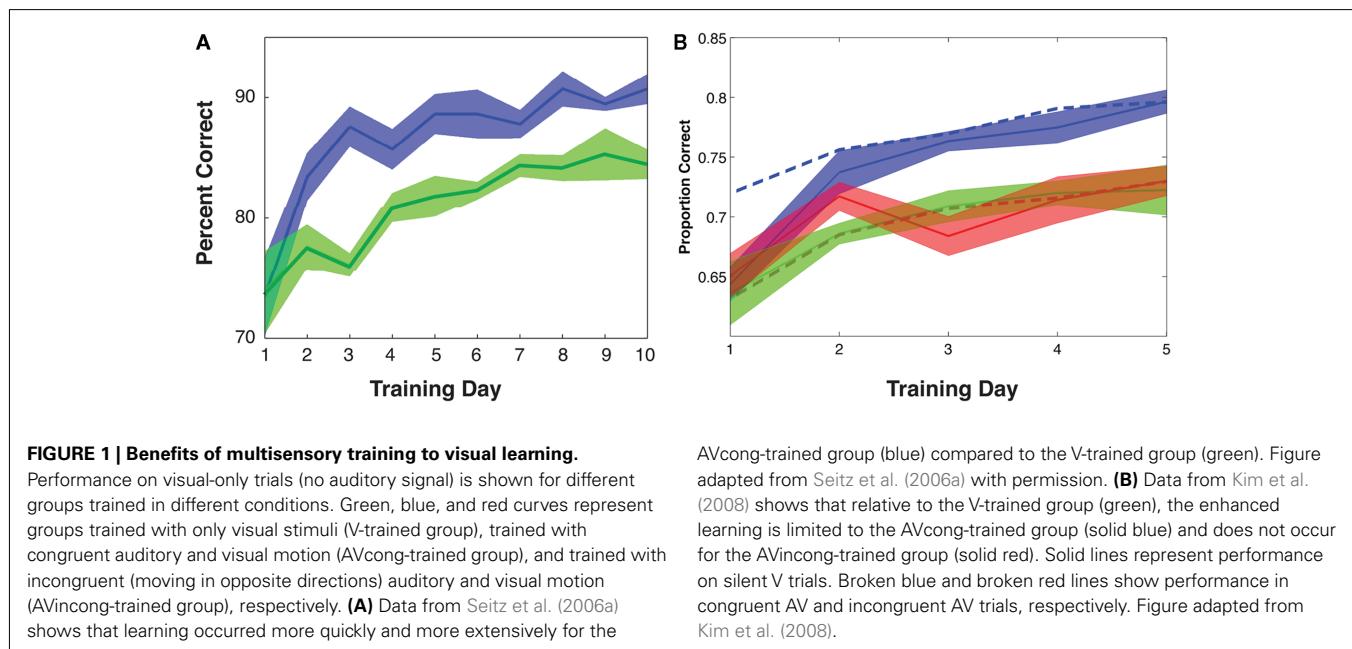
this crossmodal sensory recalibration is continuously engaged in updating unisensory perceptual processing and is an integral part of perceptual processing. In Section “Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning,” we present results from an adaptation study that shows that passive exposure to consistently paired auditory and visual features enhances visual sensitivity. These three sets of findings involve very different types of learning – perceptual learning, recalibration, and associative learning – and may involve different mechanisms and time scales, yet they all show a significant influence of multisensory processing on unisensory representations. This diversity of phenomena suggests that these multisensory influences on unisensory learning may reflect a general strategy of learning in the brain.

IMPROVEMENT IN UNISENSORY SENSITIVITY AS A RESULT OF CORRELATED MULTISENSORY TRAINING

Multisensory stimulation is widely thought to be advantageous for learning (Montessori, 1912; Fernald and Keller, 1921; Orton, 1928; Strauss and Lehtinen, 1947). As such, numerous educational programs, including the Montessori (1912, 1967) and Multisensory Structural Language Education method (Birsh, 1999), incorporate multisensory training techniques in their teaching. The benefits of multisensory training go beyond the simultaneous engagement of individuals with different learning styles (e.g., “visual learners” and “auditory learners”; Coffield et al., 2004). However, benefits of multisensory training are typically stated in anecdotal terms, such

as Treichler's (1967) statement that "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." (but see Thompson and Paivio, 1994). While the benefits of multisensory training have long been appreciated and exploited by educational and clinical practitioners, until recently there has been little solid scientific evidence to support this view.

To address the extent to which multisensory training shows benefits over unisensory training, we recently investigated how visual perceptual learning of motion–direction perception (Ball and Sekuler, 1982, 1987; Liu, 1999; Seitz et al., 2006a,b; Chalk et al., 2010; Pilly et al., 2010) is influenced by the addition of auditory information (Seitz et al., 2006a; Kim et al., 2008). Perceptual learning is an appropriate method to address the benefits of multisensory training since it is a well-established learning paradigm and a great deal is known regarding the mechanisms involved (Gilbert et al., 2001; Fahle and Poggio, 2002; Ahissar and Hochstein, 2004; Ghose, 2004; Seitz and Dinse, 2007; Shams and Seitz, 2008). We compared the effects of congruent auditory–visual (AVcong-trained) and visual (V-trained) training on perceptual learning using a coherent motion detection and discrimination task (Seitz et al., 2006a). The individuals in the AVcong-trained group were trained using auditory and visual stimuli moving in the same direction, whereas the V-trained group was trained only with visual motion stimuli. Critically, the two groups were compared on trials *without* informative auditory signals (stationary sound, and in a subsequent study described below, the two groups were compared on identical trials with no sound). Compared to the V-trained group, the AVcong-trained group showed greater learning both within the first session and across the 10 training sessions (**Figure 1A**). Therefore, *multisensory training facilitated unisensory learning*. The advantage of AV training over visual-alone training was substantial: it reduced the number of sessions required to reach asymptote by ~60%, while also raising the maximum performance.



A second study (Kim et al., 2008) showed that benefits of multisensory training were specific to training with congruent auditory–visual stimuli (i.e., moving in the same direction); a group trained with sound moving in the opposite direction of visual motion (AVincong-trained group) did not show any facilitation of learning (**Figure 1B**). This indicates that the facilitation of learning is not due to a putative alerting effect of sound during training. Additionally, results of a direction test showed that performance was significantly greater for trained directions (10° and 190°) than for untrained directions, confirming that this improvement reflects perceptual learning rather than general task learning (Ball and Sekuler, 1982; Fahle, 2004). Intriguingly, for the AVcong-trained group, the performance on silent visual trials (**Figure 1B**, solid blue) converged to the level of performance on congruent AV trials (**Figure 1B**, broken blue). In other words, individuals trained with congruent AV stimuli not only showed facilitated visual performance *when auditory stimuli were not present*, but also they performed in the absence of sound *as well as* they would perform in the presence of sound.

Other studies demonstrate that these beneficial effects are not limited to visual perceptual learning. For example, individuals trained with faces and voices can better recognize voices (auditory-alone) than those trained with voices alone (Von Kriegstein and Giraud, 2006). Memory research suggests that multisensory encoding of objects facilitates the subsequent retrieval of unisensory information (Murray et al., 2004, 2005; Lehmann and Murray, 2005). In addition, multisensory exposure has been reported to enhance unisensory reinforcement learning (Guo and Guo, 2005) in *Drosophila* (fruit flies). Collectively these studies indicate that crossmodal facilitation of learning is a general phenomenon occurring in different tasks, and across different modalities, and even species.

In a recent review, Shams and Seitz (2008) discussed how multisensory training could benefit later performance of unisensory tasks. It was suggested that facilitation could arise through two

classes of mechanisms. One possibility is that facilitation benefits learning in the same representations that undergo modification in classic unisensory learning (Seitz and Dinse, 2007). Alternatively, facilitation can be explained through multisensory exposure resulting in alterations to multisensory representations that can then be invoked by a unisensory component (Rao and Ballard, 1999; Friston, 2005). While, the findings discussed in this section can be explained by either, or a combination, of these mechanisms, other findings discussed below are suggestive that the latter mechanism (unisensory representations becoming equivalent to multisensory representations) likely play some role in the observed facilitation of learning.

CHANGE IN UNISENSORY MAP AS A RESULT OF EXPOSURE TO CROSSMODAL ERROR

As highlighted in the introduction, being endowed with multiple sensory modalities has its advantages in immediate perceptual processing. However, as illustrated in the previous section, multisensory stimulation also has a lasting effect on subsequent unisensory stimulation. This section describes the phenomenon of crossmodal sensory recalibration. Perception can generally be considered an unsupervised inference process, where the ground truth (i.e., the environmental state) is unknown, and can only be estimated from the sensorium. Therefore, comparing sensory estimates across modalities over time allows the system to perform self-maintenance by recalibrating its unisensory processes (King, 2009; Recanzone, 2009). Such changes are necessary when coping with endogenous changes that occur during development or injury, or exogenous changes in environmental conditions. An example of crossmodal recalibration is the rubber-hand illusion in which a brief (seconds) tactile stimulation of one's occluded arm while seeing a synchronous tactile stimulation of a rubber-hand subsequently induces a shift in the proprioception of the hand in the direction of the seen rubber hand (Botvinich and Cohen, 1998). Another extensively studied example of crossmodal recalibration is the ventriloquist aftereffect (VAE): the shift in perceived location of sounds (in isolation) that occurs after repeated exposure to consistent spatial discrepancy between auditory and visual stimuli (Canon, 1970; Radeau and Bertelson, 1974; Recanzone, 1998; Lewald, 2002).

While the rubber-hand illusion shows that recalibration of proprioception can occur rapidly, after seconds of exposure to tactile–visual discrepancy, recalibration of other sensory modalities such as hearing and vision has been shown to occur only after substantial exposure to spatial inconsistencies between the sensory signals, for example, after hundreds or thousands of repeated exposures to consistent discrepancy between the senses (Radeau and Bertelson, 1974; Zwiers et al., 2003; Navarra et al., 2009). In some cases, auditory recalibration has been reported after weeks, days, or hours of exposure to inconsistency (Hofma et al., 1998; Zwiers et al., 2003). The VAE has been reported to occur after several minutes of continuous exposure, or after thousands or hundreds of trials (Canon, 1970; Radeau and Bertelson, 1974; Recanzone, 1998; Lewald, 2002; Frissen et al., 2003). Altogether these results have given the impression that the human auditory and visual systems require a substantial amount of evidence that the sense is faulty before recalibration occurs.

Wozny and Shams (2011) recently conducted a study that demonstrated that auditory–visual spatial recalibration occurs much more quickly than previously thought. Observers were presented with small white disks on a black screen and white noise bursts at variable locations along azimuth for 35 ms, and were asked to localize the stimuli using a trackball that controlled the position of a cursor on the screen. On some trials only an auditory stimulus was presented, on some trials only a visual stimulus was presented, and on some trials both were presented. On bisensory trials, the observers were asked to report the location of both the visual stimulus and the auditory stimulus. All combinations of visual and auditory visual locations were presented with equal probability on both unisensory and bisensory trials, and the trials were interleaved pseudorandomly. Therefore, an auditory-alone trial could be preceded by a visual, auditory, or auditory–visual trial, and the spatial discrepancy between the auditory and visual stimuli could vary from trial to trial. This experimental design allowed us to investigate whether there is a systematic influence of AV spatial discrepancy experienced on a bisensory trial on the subsequent perception of location of sound on a unisensory auditory trial. In **Figure 2**, the change in perceived location of sound is plotted as a function of AV discrepancy in the immediately preceding AV trial. As can be seen, the perceived location of sound is shifted to the right if the auditory trial is preceded by a trial in which vision is to the right of sound, and the perceived location of sound is shifted to the left if the auditory trials is preceded by a trial in which visual stimulus was to the left of the auditory stimulus. The shift in perceived location is calculated as a difference between the reported location on a given auditory trial as compared to the reported location of sound averaged across all unisensory auditory trials with sound presented at the same location. The same qualitative results are obtained if change in perceived location is measured relative to the actual location of sound.

These findings show that auditory recalibration can occur very rapidly, after only *milliseconds* of exposure to sensory discrepancy and suggest that any exposure to discrepant auditory–visual

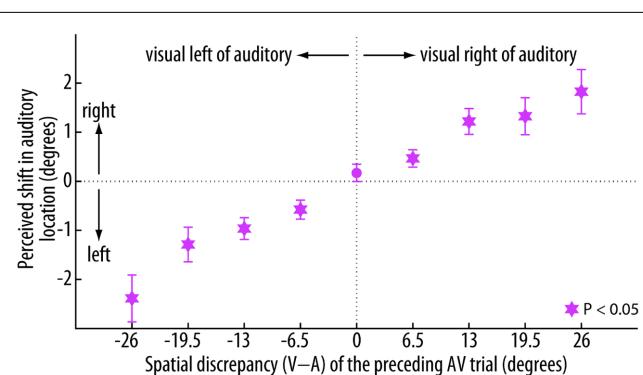


FIGURE 2 | Shift in auditory map as a function of specific exposures in the preceding trial. The shift in perceived auditory location (mean \pm SEM across observers) as a function of auditory–visual spatial discrepancy in the preceding AV trial. Stars denote datapoints that are significantly different from zero (corrected for multiple comparisons using Bonferroni–Holm correction). Figure reproduced from Wozny and Shams (2011) with permission.

sensations can instantaneously change the subsequent perception of location of sounds. This indicates a much stronger degree of malleability in our basic auditory representations (such as space) than previously thought.

Interestingly, the degree of recalibration appears to depend more on the perceived discrepancy between the auditory and visual stimuli than the physical discrepancy. The amount of recalibration was four times larger for trials in which the auditory and visual stimuli were perceived to originate from the same location than in trials where they appeared to stem from different locations (Wozny and Shams, 2011). Considering that it is not clear how long lasting the observed shifts in the auditory map are, it is possible that the recalibration phenomenon discussed here and the learning effects discussed in the previous section are mediated by distinct neural mechanisms. Studies in barn owls have found that audio–visual recalibration can involve plasticity in traditionally considered unisensory auditory and visual brain areas such as inferior colliculus (Feldman and Knudsen, 1997) and optic tectum (DeBello and Knudsen, 2004). Whether the rapid human spatial recalibration observed in Wozny and Shams (2011) involves similar mechanisms is a target of future research.

IMPROVEMENT IN UNISENSORY SENSITIVITY AS A RESULT OF MULTISENSORY ASSOCIATIVE LEARNING

While the studies described above detail how unisensory representations are altered through multisensory experience, they do not directly address how the unisensory processing is impacted by the presence of the multisensory stimulation. In a recent study, Wozny et al. (2008) investigated¹ whether after exposure to arbitrarily paired auditory and visual features, the processing of the visual feature is enhanced by the mere accompaniment of the associated auditory feature even when auditory signals are not informative for the task. If the learning of auditory–visual associations occurs at a sensory level, one could expect that the mere presence of the associated auditory feature could improve the representation of

¹These results were presented at the 2008 Vision Sciences Society Meeting and an abstract of the study is published in Journal of Vision as cited in the text.

the visual feature, however if the association is not established or if it is established at a higher level of processing, then the presence of task-irrelevant auditory signal would not enhance the visual performance (detection, discrimination, etc.). To address this issue, two experiments were conducted in which observers were passively exposed to a paired auditory–visual stimulus. In both experiments, observers demonstrated a relative increase in sensitivity to that visual stimulus when it was accompanied by the auditory stimulus that was coupled with it during exposure, even though auditory stimulus was uninformative to the subjects' task. These results suggest that unisensory benefits occur, at least in part, due to an alteration, or formation, of multisensory representations of the stimuli, as discussed in Shams and Seitz (2008).

In one experiment, oriented sinusoidal gratings were paired with pure tones. During the exposure phase, a sinusoidal grating of given visual angle of orientation (V1) was consistently presented with an auditory tone (A1) while the orthogonal orientation (V2) was presented in silence (**Figure 3A**). The visual and auditory stimuli (V1A1) co-varied in randomly chosen suprathreshold stimulus intensities across trials. The task was to keep fixation and detect any changes in the color of the fixation cross by pressing the spacebar. A change in fixation cross color occurred in approximately 10% of trials. Testing occurred prior to and after exposure. During test sessions subjects had to detect in which of two intervals the oriented grating appeared (embedded in visual noise). In trial types that involved the presentation of tones, the tone was played in both intervals and therefore, was uninformative for the task. Each test session consisted of 192 randomly interleaved trials (48 per condition). Subjects who scored close to chance (below 60%) on one or more of the pre-test conditions were excluded from sample. The two exposure conditions and four testing conditions are shown in **Figure 3A**.

A two-way repeated measures ANOVA with factors Test (pre and post) and Condition (V1A1, V1A2, V1, and V2A1) showed a significant interaction between Test and Condition [$F(3,114) = 3.86, p < 0.05$]. To determine whether passive exposure to a specific pair of auditory and visual stimuli would result in a relative increase in detection performance for that visual

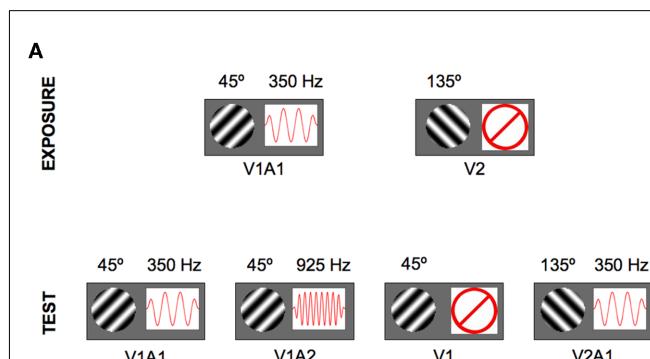
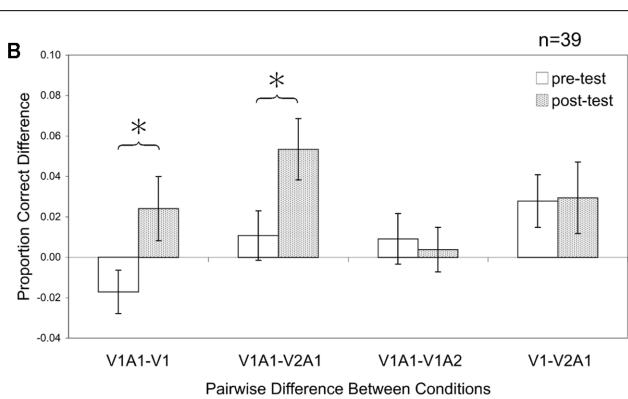


FIGURE 3 | Influence of exposure to paired visual orientation and auditory frequency on subsequent visual orientation detection. (A) Top, The stimulus conditions to which the subjects were passively exposed. Bottom, the stimulus conditions in which subjects were tested in a 2IFC



detection task. **(B)** The difference in performance between conditions before and after exposure. Stars denote significant one-tailed paired *t*-tests ($p < 0.05$) between pre and post tests corrected for multiple comparisons using the Bonferroni–Holm method. Error bars represent SE.

stimulus when accompanied by the associated sound, we compared performance differences between the pre-test and post-test data between V1A1 and V1, and found that there was a significant difference between these conditions ($p = 0.013$, one-tailed paired t -test, $df = 38$, Bonferroni–Holm $\alpha = 0.017$; **Figure 3B** column 1). If the pairing with sound had only facilitated the visual learning, the relative performance between these two conditions should have been the same. In contrast, our results suggest that an auditory–visual association was learned.

To determine whether the benefit for the V1A1 condition is a specific effect to this associated auditory–visual stimulus or whether it is a generalized effect, we examined the performance on the other testing conditions. First, if the improved performance in V1A1 is due to an alerting effect of sound, then we would expect to see the same degree of improvement in both V1A1 and V2A1. However, this was not the case as the comparison between V1A1 vs. V2A1 conditions confirmed that the facilitation was orientation specific ($p = 0.009$, one-tailed paired t -test, $df = 38$, Bonferroni–Holm $\alpha = 0.0125$; **Figure 3B** column 2). However, a significant difference was not found between learning for the exposed V1A1 condition (350 Hz tone) vs. the same orientation paired with a slightly different tone V1A2 (925 Hz), suggesting that the learning transfers across at least some range of frequencies (**Figure 3B** column 3). This degree of transfer is not entirely surprising given that the frequencies of A1 and A2 lie within an octave and a half of each other, which is within the range of auditory recalibration transfer shown in other studies (Frissen et al., 2003). Future experiments should investigate whether a wider frequency range would still show transfer of learning. Finally, as a control, we compared two conditions that had an equal amount of exposure to their components, but arranged in opponent pairings (V1 vs. V2A1) and found there was no noticeable difference in relative performance across these conditions (**Figure 3B** column 4). Altogether, these results suggest that a specific auditory–visual association was learned between V1 and A1 by passive exposure.

In the experiment described above, the auditory–visual pairing presented to subjects during exposure (V1A1) showed the greatest degree of relative improvement. This condition also happened

to be the only condition tested in which the visual stimulus was presented in the same context as that of the exposure phase. Therefore a similarity in context can be an alternative explanation for the pattern of results found in the first experiment. To address this potential confound, and to see if the effect can be replicated with other visual features, we conducted a second experiment. In this experiment, the oriented gratings were replaced by coherent dot motion. The exposure phase was similar to the first experiment, where an auditory tone (A1) was consistently paired with a particular direction of coherent motion (V1), while the orthogonal motion–direction (V2) was presented in silence. During testing, subjects had to determine the direction of coherent motion, presented with and without A1. Schematic depiction of the design is shown in **Figure 4A**, which shows the testing and exposure pairings. In contrast to the first experiment, here in addition to testing the exposed auditory visual pair V1A1, we tested V2, in which the other visual feature (not coupled with sound) is also presented in the same context (no sound) as that of the exposure phase. If the improved performance in V1A1 observed in the first experiment was due to familiar context, then similar improvement should be observed here for V2 (no-sound context). But if the improved performance was due to acquisition of a compound AV feature, then the improvement should only be observed for V1A1 and not for V2.

The exposure phase was very similar to that of Experiment 1. Subjects were presented with two trial types: V1A1 and V2. Four hundred trials of each condition were presented in pseudo-random order. Subjects were instructed to maintain fixation and to report any changes in the contrast of the fixation dot. Exposure was preceded by 256 test trials, and followed by 128 randomly interleaved test trials, 400 more exposure trials, and 128 test trials. This top-up design was used to minimize the erosion of learning effect during post-test trials. The post-test results shown below reflect the data from all 256 post-exposure trials. The entire experiment lasted about an hour. For the test sessions, a two-alternative-forced-choice (2AFC) procedure was used where a single trial was presented and the subjects were asked to report by keypress whether the coherent motion moved at 45° or 135°. Four stimulus

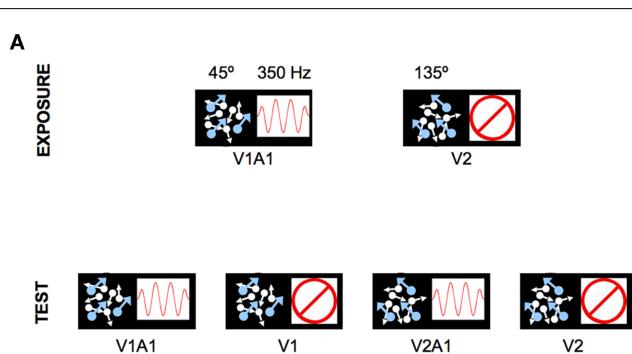
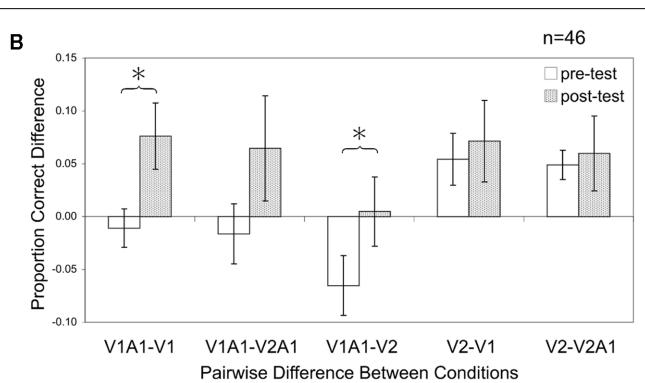


FIGURE 4 | Influence of exposure to paired visual motion–direction and auditory frequency on subsequent visual motion detection. **(A)** Top, The stimulus conditions to which the subjects were passively exposed. Bottom, the stimulus conditions in which subjects were tested in a 2IFC detection



task. **(B)** The difference in performance between conditions before and after exposure. Stars denote significant one-tailed paired t -tests ($p < 0.05$) between pre and post tests corrected for multiple comparisons using the Bonferroni–Holm method. Error bars represent SE.

conditions were tested: V1A1, V1, V2A1, and V2. Therefore, sound was not informative for the task.

This experiment replicated the findings of the first experiment. Similar to the previous experiment, we performed two-way repeated measures ANOVA with Test (pre, post) and Condition (V1A1, V1, V2A1, V2) as factors. We found a significant interaction [$F(3,135) = 2.68, p < 0.05$]. Here too, there was a significant difference between conditions V1A1 and V1 ($p = 0.007$, one-tail paired t -test, $df = 45$, Bonferroni–Holm $\alpha = 0.01$; **Figure 4B** column 1). This effect seems to be direction specific given that there is a trend of increased performance in the V1A1 conditions compared to V2A1 condition ($p = 0.036$, one-tailed paired t -test, $df = 45$, Bonferroni–Holm $\alpha = 0.0167$; **Figure 4B** column 2). The fact that the results hold true for a discrimination task in addition to the detection task used in the first experiment demonstrates that these effects are not a task-specific oddity. The fact that the V1A1 association is found for motion–direction stimuli in addition to static oriented gratings suggests that these automatic associations that we observe between the auditory and visual stimuli are a general visual phenomenon.

Another goal of this experiment was to test the hypothesis that the learning was simply due to shared context with the exposure, rather than an effect that depended on multisensory stimulation. To address this question we compared the performance between the two tested contexts that were maintained from the exposure (i.e., V1A1 and V2). We found that performance improvement from pre-test to post-test in V1A1 was superior compared to performance improvement in V2 ($p = 0.012$, one-tail paired t -test, $df = 45$, Bonferroni–Holm $\alpha = 0.0125$; **Figure 4B**, column 3). Likewise, columns 4 and 5 of **Figure 4B** show comparison V2 vs. V1 and V2 vs. V2A1, respectively. There was not any significant difference between these conditions, even though the V2 condition was equally exposed as the V1A1 condition. These results confirm that the presentation in familiar context is not the underlying factor behind the observed improvements for V1A1.

A key question is whether the exposure period creates a response bias or leads to a change in sensitivity to the stimulus. In the first experiment, we used a 2IFC paradigm in which response bias has no impact on the results. In the second experiment, we found an increase in sensitivity for the AV trials after exposure (**Figure 5A**) and no change in the bias measurements (**Figure 5B**). Our 2IFC design for the first experiment and signal detection analysis for the second experiment indicate that the improved relative performance observed for the detection/discrimination of the sound-coupled visual feature is due to an increase in sensitivity. This finding in turn suggests that the improved performance reflects learning of a low-level perceptual association. These results therefore suggest that new auditory–visual perceptual associations can be acquired based on brief exposure to correlated auditory and visual coincidences even in adult sensory systems. This indicates an impressive degree of plasticity across modalities in early sensory processing.

In contrast to previous studies of crossmodal associative learning, our study compares the effect of crossmodal associative learning on sensitivity to a visual feature with that of an exposure to the visual stimulus alone. The fact that improvement in V1A1 condition was superior to that of V2 – despite the equal exposure of V1

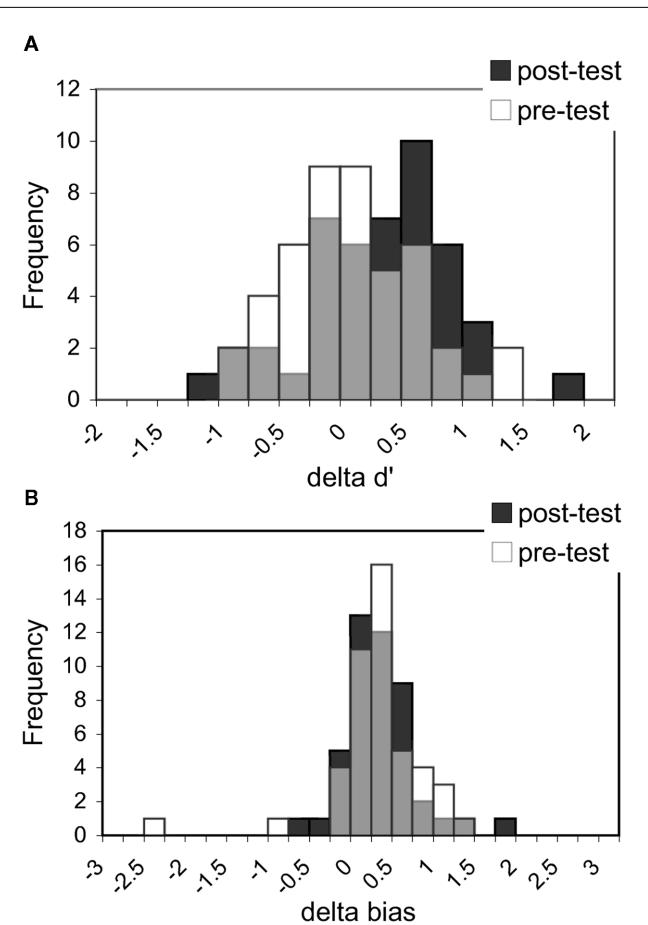


FIGURE 5 | Signal detection analysis of the experiment on associative learning of visual motion and auditory frequency. (A) A histogram of d' differences between the auditory–visual (AV) and the vision alone (V) conditions. White and dark bars show the pre- and post-test AV–V d' frequencies, respectively. Light gray bars show an overlap in distributions. For each subject, the d' values of the vision alone trials were subtracted from that of the auditory–visual trials. Thus, a positive shift in the distribution indicates an increase in sensitivity for the AV trials. **(B)** Histogram showing frequencies of bias measurements for the AV–V conditions.

and V2 – indicates that the increase in sensitivity to a visual feature achieved through establishment of a new auditory–visual feature is superior to any fine tuning of the representation obtained by exposure to the visual feature alone. This is an interesting finding, and can have important implications for perceptual skill acquisition in general. The exact mechanism by which the coupling of sound with the visual stimulus results in improved detection and discrimination of the visual stimuli is not clear. However, one possible mechanism is one in which the correlated incidence of the auditory and visual stimuli leads to establishment of new connections between the two types of feature detectors, i.e., the formation of a multisensory representation (Shams and Seitz, 2008). This will result in increased gain in the visual feature detectors whenever the visual stimulus is encountered in presence of the coupled sound. The increase in gain will in turn result in a higher sensitivity to the visual stimulus. Future studies will need to test this hypothesis.

DISCUSSION AND CONCLUSION

The human brain has evolved to learn and operate optimally in natural environments in which behavior is guided by information integrated across multiple sensory modalities. Crossmodal interactions are ubiquitous in the nervous system and occur even at early stages of perceptual processing (Shimojo and Shams, 2001; Calvert et al., 2004; Schroeder and Foxe, 2005; Ghazanfar and Schroeder, 2006; Driver and Noesselt, 2008). Until recently, however, studies of perceptual learning focused on training with one sensory modality. This unisensory training fails to tap into natural learning mechanisms that have evolved to optimize behavior in a multisensory environment.

We discussed three sets of learning phenomena that differ both in time scale and type of learning. However, in all cases multisensory exposure caused a marked change in later unisensory processing. In the learning studies discussed in Section “Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training,” the facilitation of visual learning by sound was apparent within the first hour-long session as well as across days of training. In the experiments discussed in Section “Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning,” the visual learning was evident after minutes of exposure to paired auditory–visual stimuli. The crossmodal recalibration study discussed in Section “Change in Unisensory Map as a Result of Exposure to Crossmodal Error” provided evidence that significant changes in unisensory representations can occur after only milliseconds of exposure to conflicting auditory–visual stimuli.

In the recalibration study discussed in Section “Change in Unisensory Map as a Result of Exposure to Crossmodal Error,” as well as many other previous studies of crossmodal recalibration, a mismatch between two sensory modalities (or in sensorimotor modalities) causes a change in unisensory representations. The study by Wozny and Shams (2011) shows that this adjustment

of unisensory representation based on an error signal computed from comparison with another modality does not require a protracted exposure to repeated error, and occurs continuously and incrementally. This continuous modification of unisensory representations as a result of exposure to crossmodal mismatch blurs the distinction between unisensory processing and multisensory processing. It appears that unisensory representations are closely yoked to mechanisms that keep track of crossmodal consistency/error even in the mature human nervous system.

In contrast to the learning involved in recalibration, which is caused by exposure to a mismatch between modalities, the learning phenomena discussed in Sections “Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training” and “Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning” result from exposure to multisensory stimuli that are not mismatched. In both of these cases, exposure to correlated auditory–visual stimuli causes enhanced performance in unisensory tasks. In the perceptual learning studies discussed in Section “Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training,” the multisensory stimuli are ecologically correlated, whereas in the associative learning experiments of Section “Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning” the pairing between the stimuli is arbitrary (see also Ernst, 2007). We suggest that associative and perceptual learning may represent two different stages of learning along the same dimension, with the associative learning (see Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning) representing an initial process of learning and the perceptual learning (see Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training) occurring once the association is built (see Figure 6). The idea is that initially the auditory and visual stimuli are not associated with each other in the brain,

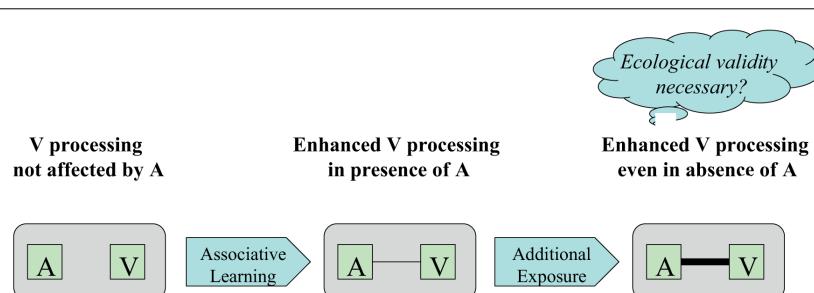


FIGURE 6 | A possible progression of learning as a result of repeated exposure to coupled auditory and visual stimuli. The representation of auditory and visual stimuli are initially not linked in the brain. Repeated exposure to paired auditory and visual stimuli results in associative learning. The newly learned association between the auditory and visual features (A and V) results in enhanced processing of the visual stimuli when accompanied by the coupled auditory stimuli. This phenomenon was discussed in Section “Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning.” For auditory and visual stimuli that are already associated in the brain, additional repeated exposure causes the connectivity/association between the two features to be strengthened further, gradually blurring the distinction between unisensory and bisensory representations (a unisensory representation becomes as effective as a

bisensory representation). This strong link between the two representations results in enhanced processing of the visual features even in the absence of the coupled auditory stimulation (and vice versa). This phenomenon was discussed in Section “Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training.” However, alternatively, the learning of association between arbitrary A and V stimuli may not progress to the phenomenon of enhanced visual processing in the absence of A. The latter phenomenon may be confined to A and V features that are ecologically related (such as motion) as it may require hard-wiring between brain areas that mediate their representations. If so, the phenomenon discussed in Sections “Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training” and “Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning” would not be parts of the same learning continuum.

and therefore the association needs to be established by repeated exposure to coupled stimuli. The establishment of the association enables the auditory stimulus to enhance the processing of the visual stimulus (and vice versa), thus improving performance in visual detection/discrimination in presence of the coupled stimulus, as described in Section “Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning.” Once this multisensory association is established, the pairing of the auditory–visual stimuli will not only improve processing at the time of stimulation (as described in Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning) but will also lead to plasticity within and between the sensory representations of these associated features, producing the facilitation and enhancement that occurs in the absence of multisensory stimulation, as described in Section “Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training.” This could be the result of visual and multisensory representations eventually becoming equivalent, where exposure to a unisensory stimulus could invoke the multisensory representation, without the need for multisensory stimulation. Such a phenomenon would result in the performance in the visual-alone and auditory–visual conditions to become equivalent, as was observed in our study (**Figure 1B**).

While we hypothesize that newly learned multisensory associations can lead to facilitation of learning, it may be the case that repeated pairing of arbitrary auditory and visual stimuli may not be sufficient to lead to lasting enhancement of unisensory processing in the absence of the crossmodal signal. It is possible that this multisensory facilitation of unisensory learning is only possible for auditory and visual features that are ecologically associated, such as auditory and visual motion, or lip movements and voice, etc. These ecologically valid associations may be distinct due to hardwired connectivity in the brain, or learning of synaptic

structures that are only possible during the critical period, and no longer possible in the mature brain. If so, then regardless of the amount of exposure, arbitrary auditory and visual features will never progress to the stage of enhanced unisensory processing in the absence of the coupled stimulus, and the phenomena discussed in Sections “Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training” and “Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning” represent two separate learning phenomena as opposed to stages of the same learning continuum. Further research is required to address these questions and to shed light on the neural and computational mechanisms mediating the three types of phenomena outlined in this paper.

We conclude that experience with multisensory stimulus arrays can have a profound impact on processing of unisensory stimuli. This can be through instant recalibrations of sensory maps (see Change in Unisensory Map as a Result of Exposure to Cross-modal Error), the formation of new linkages between auditory and visual features (see Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning), or the unisensory representations becoming increasingly indistinct from multisensory representations (see Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training). While these are operationally distinct processes, we suggest that there are linkages between the three. For example, enhancement of unisensory representations as well as recalibration of sensory maps both require establishment of their association. While further research will be required to better understand each of these types of learning, and how they relate to each other, it is now clear that the concept of unisensory processing is limited at best, and that prior multisensory exposure can affect perception within a single sensory modality even when the immediate inputs being processed are unisensory.

REFERENCES

- Ahissar, M., and Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends Cogn. Sci. (Regul. Ed.)* 8, 457–464.
- Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* 14, 257–262.
- Ball, K., and Sekuler, R. (1982). A specific and enduring improvement in visual motion discrimination. *Science* 218, 697–698.
- Ball, K., and Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Res.* 27, 953–965.
- Birsh, J. R. (1999). *Multisensory Teaching of Basic Language Skills*. Baltimore: Paul H. Brookes Publishing Co.
- Botvinich, M., and Cohen, J. (1998). Rubber hands “feel” touch that eyes see. *Nature* 391, 756.
- Calvert, G., Spence, C., and Stein, B. E. (2004). *The Handbook of Multisensory Processes*. Cambridge: The MIT Press.
- Canon, L. K. (1970). Intermodality inconsistency of input and directed attention as determinants of the nature of adaptation. *J. Exp. Psychol.* 84, 141–147.
- Chalk, M., Seitz, A. R., and Series, P. (2010). Rapidly learned stimulus expectations alter perception of motion. *J. Vis.* 10, 1–18.
- Coffield, F., Moseley, D., Hall, E., and Ecclestone, K. (2004). *Learning Styles and Pedagogy in Post-16 Learning: A Systematic and Critical Review*. London: Learning and Skills Research Centre.
- DeBello, W. M., and Knudsen, E. I. (2004). Multiple sites of adaptive plasticity in the owl’s auditory localization pathway. *J. Neurosci.* 24, 6853–6861.
- Driver, J., and Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on “sensory-specific” brain regions, neural responses, and judgments. *Neuron* 57, 11–23.
- Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *J. Vis.* 7, 7 1–14.
- Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433.
- Fahle, M. (2004). Perceptual learning: a case for early selection. *J. Vis.* 4, 879–890.
- Fahle, M., and Poggio, T. (2002). *Perceptual Learning*. Cambridge, MA: MIT Press.
- Feldman, D. E., and Knudsen, E. I. (1997). An anatomical basis for visual calibration of the auditory space map in the barn owl’s midbrain. *J. Neurosci.* 17, 6820–6837.
- 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.
- Frissen, I., Vroomen, J., De Gelder, B., and Bertleson, P. (2003). The aftereffects of ventriloquism: are they sound-frequency specific? *Acta Psychol. (Amst.)* 113, 315–327.
- Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 815–836.
- Fujisaki, W., Shimojo, S., Kashino, M., and Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nat. Neurosci.* 7, 773–778.
- Ghazanfar, A., and Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends Cogn. Sci. (Regul. Ed.)* 10, 278–285.
- Ghose, G. M. (2004). Learning in mammalian sensory cortex. *Curr. Opin. Neurobiol.* 14, 513–518.
- Gilbert, C. D., Sigman, M., and Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron* 31, 681–697.
- Gingras, G., Rowland, B. A., and Stein, B. E. (2009). The differing impact of multisensory and unisensory integration on behavior. *J. Neurosci.* 29, 4897–4902.
- Guo, J., and Guo, A. (2005). Cross-modal interactions between olfactory and visual learning in *Drosophila*. *Science* 309, 307–310.

- Hofma, P. M., Van Riswick, J. G. A., and Van Opstal, A. J. (1998). Relearning sound localization with new ears. *Nat. Neurosci.* 1, 417–421.
- Kim, R. S., Seitz, A. R., and Shams, L. (2008). Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS ONE* 3, e1532. doi:10.1371/journal.pone.0001532
- King, A. (2009). Visual influences on auditory spatial learning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 331–339.
- Lehmann, S., and Murray, M. M. (2005). The role of multisensory memories in unisensory object discrimination. *Brain Res. Cogn. Brain Res.* 24, 326–334.
- Lewald, J. (2002). Rapid adaptation to auditory-visual spatial disparity. *Learn. Mem.* 9, 268–278.
- Liu, Z. (1999). Perceptual learning in motion discrimination that generalizes across motion directions. *Proc. Natl. Acad. Sci. U.S.A.* 96, 14085–14087.
- Miyazaki, M., Yamamoto, S., Uchida, S., and Kitazawa, S. (2006). Bayesian calibration of simultaneity in tactile temporal order judgment. *Nat. Neurosci.* 9, 875–877.
- Montessori, M. (1912). *The Montessori Method*. New York: Frederick Stokes.
- Montessori, M. (1967). *The Absorbant Mind*. New York: Holt, Rinehart and Winston.
- Murray, M. M., Foxe, J. J., and Wylie, G. R. (2005). The brain uses single-trial multisensory memories to discriminate without awareness. *Neuroimage* 27, 473–478.
- Murray, M. M., Michel, C. M., Grave De Peralta, R., Ortigue, S., Brunet, D., Gonzalez Andino, S., and Schnider, A. (2004). Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *Neuroimage* 21, 125–135.
- Navarra, J., Hartcher-O'Brien, J., Piazza, E., and Spence, C. (2009). Adaptation to audiovisual asynchrony modulates the speeded detection of sound. *Proc. Natl. Acad. Sci. U.S.A.* 106, 9169–9173.
- Newell, F. N., Ernst, M. O., Tjan, B. S., and Bültthoff, H. H. (2001). Viewpoint dependence in visual and haptic objects recognition. *Psychol. Sci.* 12, 37–42.
- Orton, S. T. (1928). Specific reading disability – strephosymbolia. *JAMA* 90, 1095–1099.
- Pilly, P. K., Grossberg, S., and Seitz, A. R. (2010). Low-level sensory plasticity during task-irrelevant perceptual learning: evidence from conventional and double training procedures. *Vision Res.* 50, 424–432.
- Radeau, M., and Bertelson, P. (1974). The after-effects of ventriloquism. *Q. J. Exp. Psychol.* 26, 63–71.
- 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.
- Recanzone, G. (1998). Rapidly induced auditory plasticity: the ventriloquism aftereffect. *Proc. Natl. Acad. Sci. U.S.A.* 95, 869–875.
- Recanzone, G. H. (2009). Interactions of auditory and visual stimuli in space and time. *Hear. Res.* 258, 89–99.
- Schroeder, C. E., and Foxe, J. (2005). Multisensory contributions to low-level, “unisensory” processing. *Curr. Opin. Neurobiol.* 15, 454–458.
- Seitz, A. R., and Dinse, H. R. (2007). A common framework for perceptual learning. *Curr. Opin. Neurobiol.* 17, 148–153.
- Seitz, A. R., Kim, R., and Shams, L. (2006a). Sound facilitates visual learning. *Curr. Biol.* 16, 1422–1427.
- Seitz, A. R., Nanez, J. E., Holloway, S. R., Tsushima, Y., and Watanabe, T. (2006b). Two cases requiring external reinforcement in perceptual learning. *J. Vis.* 6, 966–973.
- Shams, L., and Seitz, A. (2008). Benefits of multisensory learning. *Trends Cogn. Sci. (Regul. Ed.)* 12, 411–417.
- Shimojo, S., and Shams, L. (2001). Sensory modalities are not separate modalities: plasticity and interactions. *Curr. Opin. Neurobiol.* 11, 505–509.
- Strauss, A., and Lehtinen, L. E. (1947). *Psychopathology and Education of the Brain-Injured Child*. New York: Grune & Stratton.
- Sumby, W. H., and Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.* 26, 212–215.
- Thompson, V., and Paivio, A. (1994). Memory for pictures and sounds: independence of auditory and visual codes. *Can. J. Exp. Psychol.* 48, 380–396.
- Treichler, T. G. (1967). Are you missing the boat in training aid? *Film AV Commun.* 1, 14–16.
- Von Kriegstein, K., and Giraud, A. L. (2006). Implicit multisensory associations influence voice recognition. *PLoS Biol.* 4, e326. doi:10.1371/journal.pbio.0040326
- Wozny, D. R., Seitz, A. R., and Shams, L. (2008). Learning associations between simple visual and auditory features. *J. Vis.* 8, 171.
- Wozny, D. R., and Shams, L. (2011). Recalibration of auditory space following milliseconds of cross-modal discrepancy. *J. Neurosci.* 31, 4607–4612.
- Zwiers, M. P., Van Opstal, A. J., and Paige, G. D. (2003). Plasticity in human sound localization induced by compressed spatial vision. *Nat. Neurosci.* 6, 175–181.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 09 July 2011; accepted: 23 September 2011; published online: 18 October 2011.

Citation: Shams L, Wozny DR, Kim R and Seitz A (2011) Influences of multisensory experience on subsequent unisensory processing. Front. Psychology 2:264. doi: 10.3389/fpsyg.2011.00264

This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.

Copyright © 2011 Shams, Wozny, Kim and Seitz. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.