Correspondence

Visual rhythm perception improves through auditory but not visual training

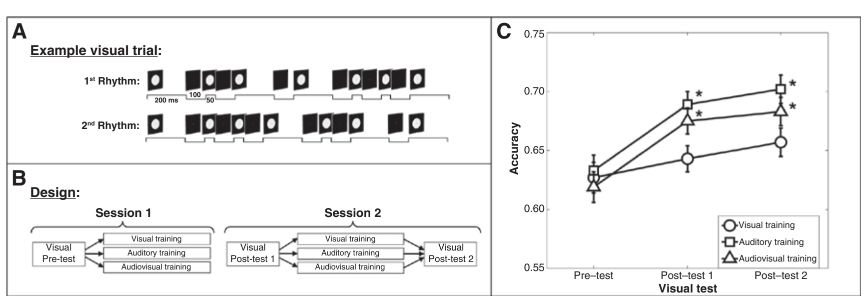
Brandon Barakat[1], Aaron R. Seitz[2], and Ladan Shams[3]

Memory research has shown that test performance is optimal when testing and practice occur in identical contexts [1]. However, recent research in object recognition and perceptual learning has shown that multisensory practice leads to improved test performance, even when the test is unisensory [2,3]. It is also known that different sensory modalities can have differing proficiencies in a given domain. For instance, research shows that, compared to the auditory modality, the visual modality is significantly less proficient at discriminating the rhythms of temporal sequences [4,5]. Although rhythm perception is typically thought of as residing in the auditory domain, instances of visual rhythm perception abound in daily life, for example, when one watches a dancer or a drummer, or when a doctor examines a patient's breathing or heart rate on a monitor (such as when diagnosing arrhythmia). However, no previous study has examined whether visual rhythm discrimination is a trainable perceptual skill. In light of this, we examined the extent to which visual rhythm perception can be improved through two sessions of visual, auditory, or audiovisual training. We found that visual rhythm discrimination was significantly improved in the auditory and audiovisual training groups, but not in the visual training group. Our results show that, in certain tasks, within-modality training may not be the best approach and that, instead, training in a different sensory modality can be a necessary approach to achieve learning.

Using a sample size determined by a power analysis of pilot data, 30 subjects were randomly assigned to each of the three training conditions: visual (blinking disc), auditory (pure-tone beeps), or audiovisual (temporally congruent blinking disc and beeps). The stimuli and task were adapted from Saenz and Koch [4]. Across two sessions, each group was trained and tested in a rhythm discrimination task, wherein they judged whether two rhythm sequences, made of an assortment of short and long stimulus pulses, were the same or different. Whereas the stimulus modality used in the training trials differed across groups, the test trials consisted of only visual stimuli for all groups. On each trial, subjects were presented with two successive rhythm sequences, separated by a 1.5 s pause. Each rhythm sequence consisted of seven elements, including three short (50 ms) and four long (200 ms) duration elements. The inter-stimulus interval between the elements was 100 ms. A standard rhythm sequence was uniquely generated for each subject by pseudo-randomizing the serial position of the seven elements and this standard rhythm was then used throughout training and testing. See Figure 1A for a schematic of an example visual trial wherein the two rhythms are different.

In the first session, subjects completed 10 practice trials (visual stimuli only), one test block (pre-test; 50 trials) and four blocks of training (50 trials per block). In the second session, subjects completed 10 warm-up trials (visual stimuli only), a test block (post-test 1; 50 trials), three blocks of training (50 trials per block), and a final test block (post-test 2; 50 trials). To track subjects' visual learning during training, 10 visual-only test trials were randomly interleaved into each training block for all groups. A depiction of the experimental design can be seen in Figure 1B. Subjects received feedback during practice and training, but no feedback was given during tests. At the end of the second session, subjects completed an exit questionnaire in which they reported their number of years of formal music training. This measure was included as a control variable, as past research has shown that musicians have significantly better visual and auditory rhythm perception compared to non-musicians [6]. Additional details about our experimental procedures can be found in the Supplemental Information available online.

An analysis of the questionnaire responses revealed that years of music training did not differ across the three groups (see Supplemental Information for more details) and that there was a significant positive correlation between years of music training and accuracy in the pre-test ($r = 0.32$, $p = 0.002$). In light of this, subjects' percent accuracy in the visual rhythm discrimination task was analyzed using an ANCOVA with Test (pre-test, post-test 1, post-test 2) as an independent variable and years of music training entered as a covariate, for each of the three groups. Despite two sessions of training, the visual training group failed to experience significant improvement in their visual rhythm perception ($p = 0.42$, $\eta_p^2 = 0.02$). In



Current Biology

Figure 1. Experimental procedures and results of visual rhythm discrimination training.
(A) Schematic of a visual trial from the rhythm discrimination task. In this example, the patterns of the rhythm sequences are different. (B) Depiction of the experimental design. Sessions were completed 24 to 48 hours apart. For all groups, the pre- and post-test trials contained visual stimuli only. (C) Mean percent accuracy in each visual test is depicted for each training condition. Asterisks indicate a significant change from pre-test ($p \leq 0.01$). Error bars represent within-subject standard error of the mean.

contrast, visual rhythm perception was significantly improved in the groups that received auditory (p = 0.007, η_p^2 = 0.11) and audiovisual (p = 0.01, η_p^2 = 0.10) training. This pattern of results can be seen in Figure 1C. An analysis of reaction times indicated that the observed changes in accuracy were not due to a speed–accuracy trade-off. Additional details regarding the reaction time data and training data can be found in the Supplemental Information.

To our knowledge, this is the first study to show that rhythm perception is susceptible to perceptual learning. Furthermore, this appears to be a long-term learning effect, as the post-tests were completed 24 to 48 hours after the initial training. The idea that visual perceptual learning can be improved by engaging another modality during training is not new. For instance, audiovisual training has been found to facilitate visual perceptual learning in a motion discrimination task [2], and auditory training has been shown to transfer to the visual modality in a temporal interval discrimination task [7]. However, in prior work, training in one modality and then testing in another has resulted in less learning when compared to training and testing in a single modality (for example [7]). Here we show the opposite, such that visual-only training *failed* to improve visual rhythm discrimination, whereas auditory and audiovisual training produced substantial learning. Given that past research has shown that multisensory training conditions can result in a synergistic learning effect (for example [2]), it is interesting that the audiovisual group did not show an advantage over the auditory group. A potential explanation for this is that the auditory modality was able to perform the rhythm discrimination task with such ease, and the visual modality with such difficulty, that any potential synergistic multisensory facilitation of learning failed to occur. Nonetheless, the current results demonstrate that performance in one modality can not only be facilitated by engaging another modality during training, but that in certain tasks it may be a necessary condition to produce learning when unisensory training alone is insufficient.

Our findings also have important implications for our understanding of learning and plasticity in the brain. Current debates in the field of perceptual learning revolve around relative contributions of different early vs. late stage mechanisms in supporting learning [8]. Our findings are consistent with late stage perceptual learning and enrich this conversation by showing that, in some cases, perceptual processing and perceptual learning may depend on representations that are not native to the task but that are instead ‘out-sourced’ (for example, crossmodal read-out of visual representations by systems that are also read-out auditory representations). This possibility is in line with research indicating that temporal information may be automatically encoded by the auditory modality, even when that information is delivered via the visual modality [9,10]. For instance, Kanai *et al*. [10] found that disrupting auditory cortex via transcranial magnetic stimulation (TMS) resulted in deficits in temporal judgments (that is, duration discrimination) in both the auditory and visual modalities, whereas applying TMS to visual cortex interfered with visual, but not auditory, temporal judgments [10].

Given that perceptual learning of rhythm is a largely unexplored subject, there are many questions for future studies to address. For example, it is unknown whether the observed improvements in rhythm perception were specific to the standard rhythm or if they would generalize to novel rhythm sequences. Furthermore, it is unclear from the current findings whether additional improvements in visual rhythm perception could have been attained by including more training sessions. Future studies could also examine the longevity of the observed learning effects, by administering additional tests after a delay (for example, across weeks or months). Finally, future research should explore the extent to which the current findings have relevance in clinical and educational applications, wherein rehabilitation and/or skill acquisition may require training protocols that employ sensory modalities that are more apt for performing a given task than the sensory modalities that are intended to be improved.

Supplemental Information

Supplemental information includes results, experimental procedures, two figures, and four movies, and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2014.12.011.

Acknowledgments

This research and Brandon Barakat were funded by National Science Foundation Grant BCS-1057625.

References

1. Hirshman, E., and Bjork, R.A., (1988). The generation effect: Support for a two-factor theory. J. Exp. Psychol. Learn. *14*, 484–494.
2. Seitz, A.R., Kim, R., and Shams, L. (2006). Sound facilitates visual learning. Curr. Biol. *16*, 1422–1427.
3. Lehmann, S., and Murray, M.M. (2005). The role of multisensory memories in unisensory object discrimination. Cogn. Brain Res. *24*, 326–334.
4. Saenz, M., and Koch, C. (2008). The sound of change: Visually-induced auditory synesthesia. Curr. Biol. *18,* R650–R651.
5. Glenberg, A.M., Mann, S., Altman, L., Forman, T., and Procise, S. (1989). Modality effects in the coding and reproduction of rhythms. Mem. Cogn. *17,* 373–383.
6. Rammsayer, T.H., Buttkus, F., and Altenmueller, E. (2012). Musicians do better than nonmusicians in both auditory and visual timing tasks. Music Percept. *30*, 85–96.
7. Bratzke, D., Seifried, T., and Ulrich, R. (2012). Perceptual learning in temporal discrimination: Asymmetric cross-modal transfer from audition to vision. Exp. Brain Res. *221,* 205–210.
8. Watanabe, T., and Sasaki, Y. (2014). Perceptual learning: Towards a comprehensive theory. Annu. Rev. Psychol. http://dx.doi.org/10.1146/annurev-psych-010814-015214, Sept. 10, 2014.
9. Karabanov, A., Blom, Ö., Forsman, L., and Ullén, F. (2009). The dorsal auditory pathway is involved in performance of both visual and auditory rhythms. Neuroimage *44*, 480–488.
10. Kanai, R., Lloyd, H., Bueti, D., and Walsh, V. (2011). Modality-independent role of the primary auditory cortex in time estimation. Exp. Brain Res. *209*, 465–471.

[1]Department of Psychology, University of California, Los Angeles, Los Angeles, CA 90095, USA. [2]Department of Psychology, University of California, Riverside, Riverside, CA 92521, USA. [3]Department of Psychology, University of California, Los Angeles, Los Angeles, CA 90095, USA.
E-mail: bkbarakat@ucla.edu, aseitz@ucr.edu, ladan@psych.ucla.edu