

Journal Club

Editor's Note: These short, critical reviews of recent papers in the *Journal*, written exclusively by graduate students or postdoctoral fellows, are intended to summarize the important findings of the paper and provide additional insight and commentary. For more information on the format and purpose of the Journal Club, please see http://www.jneurosci.org/misc/ifa_features.shtml.

Human Learning Improves Machine Learning: Neural and Computational Mechanisms of Perceptual Training

Joshua Carp and Joonkoo Park

Department of Psychology, University of Michigan, Ann Arbor, Michigan 48109

Review of Zhang et al.

Visual perception is highly plastic; perceptual training dramatically increases sensitivity to contrast, contour, and motion. However, many questions remain about the neural mechanisms underlying such learning-related change. Previous research shows that practice can increase neural selectivity for trained stimuli (Jiang et al., 2007). Such selectivity might reflect increased neural responses to preferred stimuli, decreased responses to nonpreferred stimuli, or a combination of the two. These issues have both theoretical and applied significance; improved understanding of the mechanisms supporting perceptual learning could lead to important advances in the explanation and treatment of visual disorders, including amblyopia and age-related visual impairment.

Zhang et al. (2010) used multivoxel pattern analysis (MVPA) to explore the neural mechanisms underlying practice-related improvement in the perception of visual form. MVPA uses fine-scale patterns of neural activation, measured using functional MRI, to classify observers' perceptual choices. This method has been used to decode edge orientation and motion direction from brain responses to visual stimuli (Kamitani and Tong, 2005). Zhang et al. (2010) measured neural activation

in a form perception task before and after psychophysical training in human subjects. Observers viewed patterns that were parametrically varied along a spectrum from radial to concentric forms (Zhang et al., 2010, their Fig. 1A). The authors then used MVPA to classify patterns of neural activity according to the parameters of the form stimulus that was presented on each trial. To assess the accuracy of these classifications, they calculated the likelihood of misclassification as a function of the similarity between the presented stimulus and the classifier's prediction.

Perceptual training improved classification performance based on activity in ventral and dorsal visual areas in two ways. First, training increased the number of stimuli classified correctly by MVPA (i.e., the amplitude of the tuning function). Zhang and colleagues (2010) attributed this result to enhanced blood oxygenation level-dependent responses across voxels that encode the preferred stimulus category. Second, training decreased the number of grossly misclassified patterns (i.e., the SD of the tuning function): when the classifier predicted visual form incorrectly, the predicted form tended to be more similar to the true form posttraining (Zhang et al., 2010, their Fig. 2). The authors attributed this result to decreased neural responses to nonpreferred stimuli. The results also extend our knowledge about the role of the dorsal visual stream in form perception. Previous research has principally linked

form perception with the ventral stream, or the "what" pathway; Zhang and colleagues (2010) show that the dorsal "where" pathway contributes to this process as well.

Zhang and colleagues (2010) showed that training improves both behavioral performance and neural selectivity in a form perception task. These intriguing results raise many questions about the mechanisms underlying perceptual learning. Successful task performance relies on a series of computational processes: observers must detect and identify stimulus input, apply decision rules, and maintain attention to the task. Which of these processes change during perceptual learning?

The training effects observed by Zhang et al. (2010) may reflect bottom-up changes in the perception of visual form. For example, perceptual learning might improve observers' ability to detect visual stimuli, to integrate contrast information across local or global scales, or to segment signal dots from noise. Alternatively, learning might affect top-down control of visual perception. Task performance relies on decision processes that compare the outputs of competing pools of sensory neurons (Heekeren et al., 2004). Observers must also maintain focus on the task; performance declines during lapses of attention (Christoff et al., 2009). Training might improve these abilities as well. For example, practice-related improvements in form perception in the present study might reflect gains in the ability to avoid mind-wandering or lapses of attention.

Received Dec. 20, 2010; revised Jan. 13, 2011; accepted Jan. 19, 2011.

Correspondence should be addressed to Joshua Carp, Department of Psychology, University of Michigan, 530 Church Street, Ann Arbor, MI 48109. E-mail: jmcarp@umich.edu.

DOI:10.1523/JNEUROSCI.6644-10.2011

Copyright © 2011 the authors 0270-6474/11/313937-02\$15.00/0

Future experiments should parse the contributions of bottom-up and top-down processes to the training effects documented by Zhang et al. (2010). Top-down mechanisms related to perceptual decision-making and attentional control are thought to rely on prefrontal representations (Heekeren et al., 2004). If learning improves top-down modulation of perception, then training might alter neural responses in prefrontal as well as visual areas. In contrast, if learning sharpens the tuning of neural-form representations via bottom-up mechanisms, then learning-related effects on neural tuning should persist in the absence of top-down control—for example, when form stimuli are presented outside the focus of attention or outside conscious awareness. Indeed, psychophysical studies show that perceptual learning can occur even for subconsciously presented stimuli (Watanabe et al., 2001), reflecting changes in bottom-up mechanisms. However, if Zhang et al.'s (2010) results principally reflect changes in top-down processes, then the benefits of training should only emerge when observers attend to form stimuli.

The present results also raise questions about the nature of the information extracted by MVPA: did the SVM classifier used in the present experiment detect representations of objective information in the visual display or subjective interpretations of that information? Serences and Boynton (2007) showed that activation patterns in both early and late visual areas classified the true direction of visual motion stimuli. In contrast, only late visual activation classified observers' subjective perception of motion independent of visual input. Future studies should extend these results to the form perception task used in Zhang et al. (2010), assessing the contribution of each experience-sensitive visual area to objective versus subjective coding.

Zhang and colleagues (2010) showed that perceptual training simultaneously enhances neural responses to preferred stimuli and suppresses responses to non-preferred stimuli. These two observations might reflect a single underlying phenomenon. For example, learning might directly increase neural responses to preferred-form stimuli. This increased activation might, in turn, decrease re-

sponses in neurons that represent non-preferred stimuli via local inhibitory circuits. Alternatively, enhancement and suppression might stem from independent mechanisms, with dissociable effects on perception and decision making. Studies of the time course of perceptual training over multiple sessions may elucidate the relationship between these phenomena. Do enhancement and suppression develop in parallel, or does one effect lag the other in time? Are the two phenomena correlated across trials, sessions, or subjects? How does each effect relate to learning-related improvements in behavioral performance?

Zhang et al.'s (2010) results also raise interesting questions about the durability and generality of training effects on behavioral performance and brain function. Observers were scanned before and after three psychophysical training sessions. Learning-related improvement in motion perception may last 10 weeks or longer beyond the end of practice (Ball and Sekuler, 1982). Neural and behavioral correlates of form-perception training may be similarly long-lived; alternatively, such effects may decay quickly in the absence of continued practice. Future studies should also determine the extent to which the effects of visual-form training generalize to other tasks. For example, if the learning-related effects observed by Zhang and colleagues (2010) reflect general improvements in form perception, this training protocol should also improve the perception of contrast- and motion-defined form, as well as the luminance-defined forms used here. Training might also affect the representation of naturalistic visual forms, including faces and scenes. In contrast, if this protocol improves form tuning by reducing mind-wandering or increasing motivation to attend to the task, training might improve performance across nearly any situation that requires observers to sustain attention to a dull or repetitive task. Thus, future studies should assess transfer effects from the present training protocol to a broad range of perceptual tasks.

Finally, these results illustrate the difficulty of interpreting the sophisticated techniques increasingly used in modern neuroimaging. The present experiment submitted the raw data to several stages of

analysis. The researchers first submitted activation estimates to 15 pairwise SVM classifiers. Next, they aggregated these two-way classifiers into a single six-way classifier. Finally, they fit Gaussian tuning curves to the output of the six-way classifier. The complexity of this procedure obfuscates the relationships between successive stages of analysis: the output of one stage may differ from that of the next in unexpected ways. For example, although the accuracy of the six-way classifier closely paralleled the amplitude of the fitted tuning function in most regions, these estimates yielded qualitatively different results in area V4v. Before fitting, classification accuracy was higher for pre-training than post-training; after fitting, this pattern was reversed (compare Zhang et al., 2010, their Figs. 3A and S3). Similarly, fitting reduced the confidence intervals of tuning amplitude by an order of magnitude relative to classification accuracy. In summation, such multilayered analytic techniques can substantially change the meaning of the data. Researchers should apply and interpret these methods with caution.

References

- Ball K, Sekuler R (1982) A specific and enduring improvement in visual motion discrimination. *Science* 218:697–698.
- Christoff K, Gordon AM, Smallwood J, Smith R, Schooler JW (2009) Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proc Natl Acad Sci U S A* 106:8719–8724.
- Heekeren HR, Marrett S, Bandettini PA, Ungerleider LG (2004) A general mechanism for perceptual decision-making in the human brain. *Nature* 431:859–862.
- Jiang X, Bradley E, Rini RA, Zeffiro T, Vanmeter J, Riesenhuber M (2007) Categorization training results in shape- and category-selective human neural plasticity. *Neuron* 53:891–903.
- Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8:679–685.
- Serences JT, Boynton GM (2007) The representation of behavioral choice for motion in human visual cortex. *J Neurosci* 27:12893–12899.
- Watanabe T, Náñez JE, Sasaki Y (2001) Perceptual learning without perception. *Nature* 413:844–848.
- Zhang J, Meeson A, Welchman AE, Kourtzi Z (2010) Learning alters the tuning of functional magnetic resonance imaging patterns for visual forms. *J Neurosci* 30:14127–14133.