Perceptual learning of oriented gratings as revealed by classification images

Jonathan Dobres
Department of Cognitive and Neural Systems, Center for Adaptive Systems, Center of Excellence for Learning in Education, Science, and Technology, Boston University, Boston, MA, USA

Aaron R. Seitz
Department of Psychology, University of California, Riverside, Riverside, CA, USA

Classification image analysis is a psychophysical technique in which noise components of stimuli are analyzed to produce an image that reveals critical features of a task. Here we use classification images to gain greater understanding of perceptual learning. To achieve reasonable classification images within a single session, we developed an efficient classification image procedure that employed designer noise and a low-dimensional stimulus space. Subjects were trained across ten sessions to detect the orientation of a grating masked in noise, with an eleventh, test, session conducted using a stimulus orthogonal to the trained stimulus. As with standard perceptual learning studies, subjects showed improvements in performance metrics of accuracy, threshold, and reaction times. The clarity of the classification images and their correlation to an ideal target also improved across training sessions in an orientation-specific manner. Furthermore, image-based analyses revealed aspects of performance that could not be observed with standard performance metrics. Subjects with threshold improvements learned to use pixels across a wider area of the image, and, apposed to subjects without threshold improvements, showed improvements in both the bright and dark parts of the image. We conclude that classification image analysis is an important complement to traditional metrics of perceptual learning.

Keywords: perceptual learning, classification images, orientation discrimination

Citation: Dobres, J., & Seitz, A. R. (2010). Perceptual learning of oriented gratings as revealed by classification images. *Journal of Vision, 10*(13):8, 8–11, http://www.journalofvision.org/content/10/13/8, doi:10.1167/10.13.8.

Introduction

Perceptual learning is defined as any relatively permanent or consistent change in an observer's perception of a stimulus following experience of that stimulus (Gibson, 1963). Perceptual learning has been shown to occur for a variety of basic visual features, such as orientation and spatial frequency of gratings (Fiorentini & Berardi, 1980) and direction of motion (Ball & Sekuler, 1982). The performance improvements that occur over the course of training tend to follow characteristic patterns that allude to aspects underlying brain plasticity (Ahissar & Hochstein, 2004; Dosher & Lu, 1998; Fahle, 2004; Seitz & Dinse, 2007). For instance, a performance improvement with a stimulus of one spatial orientation may not transfer to another orientation (Fiorentini & Berardi, 1980). It has been argued that this pattern of results is consistent with changes in orientation selectivity of early visual cells (Schoups, Vogels et al., 2001) and indicates that the locus of learning may be in these early visual areas (Fahle, 2004). However, the argument that specificity of learning demonstrates plasticity in visual cortex is not well supported (Law & Gold, 2008; Xiao, Zhang et al., 2008) and it is clear that task performance alone is an insufficient measure by which to understand brain mechanisms involved in perceptual learning.

The relationship between psychophysical performance and underlying physiological changes is fundamentally inferential. Typically, perceptual learning is operationalized as improvement in sensitivity (such as threshold or accuracy) or reaction time on a task. These metrics have provided a great deal of insight into the mechanisms of perception and perceptual learning, but the inferential gap between the observer's perception and the final performance measurement remains problematic (Mollon & Danilova, 1996). One aspect of this problem is that sensitivity and reaction time are very gross measures of performance and reveal little detail of observers' perceptual processes. As a result, it is the onus of the experimenter to design clever studies to rule out alternative explanations of changes in these performance metrics. This task is further complicated by the fact that, as we report here, these two metrics can produce opposite patterns under some conditions.

We find an alternative and potentially richer metric in classification images (Ahumada, 2002). In a classification image analysis, observers detect or discriminate a stimulus of interest (the signal) embedded in external noise. In a typical experiment, an observer is presented with a signal-with-noise stimulus and a noise-only stimulus

doi: 10.1167/10.13.8 Received April 14, 2010; published November 10, 2010 ISSN 1534-7362 © ARVO

in succession and asked to report which stimulus contained the target image. The noise fields from these stimuli are then grouped and analyzed according to the observer's decisions, ultimately producing a classification image that may be thought of as the mental template that the observer used to classify stimuli during the task. Classification images have been produced from a wide variety of studies and reveal important aspects of perception in both low- and high-level visual tasks (Keane, Lu et al., 2007; Lu & Liu, 2006; Mareschal, Dakin et al., 2006; Sekuler, Gaspar et al., 2004; Shimozaki, Chen et al., 2007). The advantages of using classification images as a metric of perceptual learning are substantial because classification images encode much more detailed characteristics of an observer's perceptual processes than do sensitivity or reaction time measures.

However, one limitation of classification image methods is that a large number of trials is required to produce a stable classification image. In many studies, upward of 10,000 trials are employed to construct a classification image (Ahumada, 2002; Lu & Liu, 2006; Sekuler et al., 2004). Such a large number of trials, which need to be spread out across multiple sessions, can be problematic to capture effects of perceptual learning because substantial learning can occur during the acquisition of the classification images. To successfully examine perceptual learning of lower order visual features over the course of several days requires the acquisition of an image in a single experimental session, approximately 1,000 trials. This has been accomplished in the case of perceptual learning of vernier acuity (Li, Levi et al., 2004), however, that study involved stimuli with only 16 changing parts, thus benefiting from a relatively low-dimensional stimulus space. Generating a classification image of a stimulus that is more extended in space (such as an oriented grating) presents a greater challenge because such images require variations in many hundreds of pixels and involve a much higher dimensional stimulus space.

In this paper, we employ an efficient classification image technique that makes it possible to produce stable images from as little as 512 trials, allowing for an image-based analysis of perceptual learning over the course of several days. We examine both traditional metrics of perceptual learning (such as reaction time, threshold, and accuracy) as well as changes in observers' classification images.

Methods

Subjects

A total of 14 subjects participated in this experiment. All had normal or corrected-to-normal vision. Subjects provided informed consent and were tested under conditions that conformed to the guidelines of the University of California, Riverside Human Research Review Board. One subject was excluded from analysis due to irregularities in compliance with experiment protocol, and one was excluded due to recurrent technical problems with the testing equipment.

Apparatus

An Apple Mac Mini running Matlab (Mathworks, Natick, MA) and Psychtoolbox Version 3 (Brainard, 1997; Pelli, 1997) was used for stimulus generation and experiment control. Subjects sat on a height adjustable chair at a distance of 36 in from a 19.5-in horizontally wide Dell Trinitron CRT monitor set to a resolution of 1024×768 and a refresh rate of 100 Hz. The distance between the subjects' eyes and the monitor was fixed by having them position their head in a chin rest with a head bar. Care was taken such that the eyes and the monitor center were at the same horizontal level.

Stimuli

The stimuli were designed such that each trial would contribute maximally to the classification image. Stimuli were 16×16 pixels in resolution, enlarged to subtend 5° of visual angle and presented centrally. Stimuli consisted of a low-resolution square-wave grating, approximately 0.8 cycles per degree, embedded in noise generated from an m-sequence (Reid, Victor et al., 1997). The use of an m-sequence resulted in pseudorandom noise images composed of black and white pixels without shades of gray and ensured that the pairwise correlations of pixels across stimuli were minimized. A total of 256 noise masks were generated from the m-sequence. This set of masks was then duplicated four times, and half of these were contrast-inverted. This ensured that the mean noise mask across all trials was the same for all pixels.

Procedure

The general procedure for the experiment is shown in Figure 1. Subjects conducted a two-interval forced-choice (2IFC) task. On each trial, one interval would display a noise mask generated from the m-sequence, while the other would display a different noise mask with a percentage of the signal stimulus (the square-wave grating) added into it via a pixel substitution method (Nishina, Seitz et al., 2007; Seitz, Kim et al., 2009). Trials were evenly split between either a constant signal level of 30% or a signal level determined by a 3-down, 1-up staircase. Each stimulus was displayed for 300 ms, with an interstimulus interval of 500 ms. Subjects were required to respond within 1,500 ms after the offset of the second

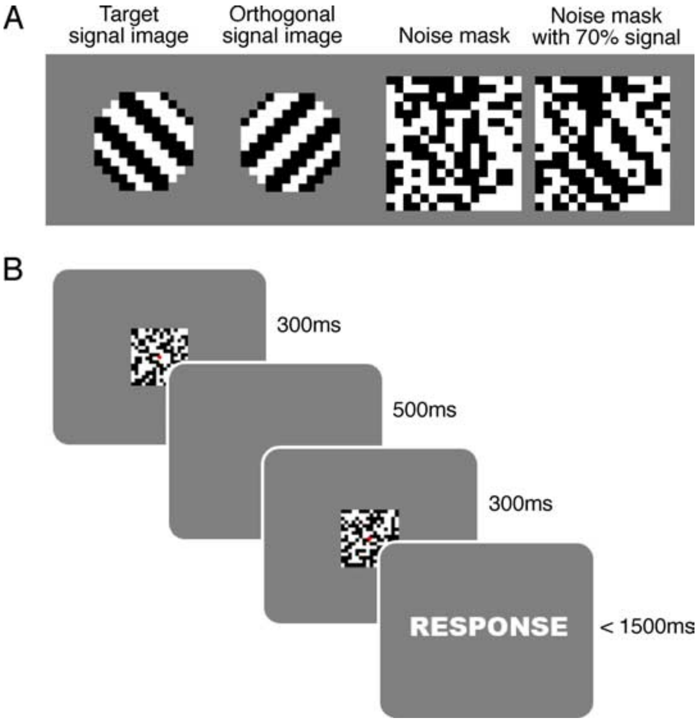

Figure 1. Schematic of the experiment design. (A) Examples of training and transfer signal images, as well as a noise mask with and without signal image added. Different masks were used during each interval of a single trial. (B) Schematic of a single trial.

stimulus and were provided trial-by-trial feedback regarding the correctness of their responses.

Each session consisted of 1,024 trials and lasted approximately 1 h. Subjects participated in a total of eleven sessions. During the first ten sessions, subjects were trained to detect a grating oriented at either 45° or 135° (orientation counterbalanced across subjects). In the eleventh session, a transfer test was conducted in which all presented stimuli were rotated to the orientation orthogonal to the one used in training. Each session was split into eight blocks. At the start of each block, subjects were shown the target signal and reminded to "keep this in mind" as he/she performed the task.

Monte Carlo filter

A Monte Carlo-based image filter was used to extract components of the classification image that were unlikely to occur by chance. Stimuli from each session were reprocessed with a random pattern of simulated answers that approximated chance (50% correct) performance. A classification image was then computed from this random simulation. The process was repeated 10,000 times for each session's stimulus set. This established a distribution of possible values for each pixel of every classification image in both Cartesian and Fourier spaces. Pixel data of classification images were then compared against these

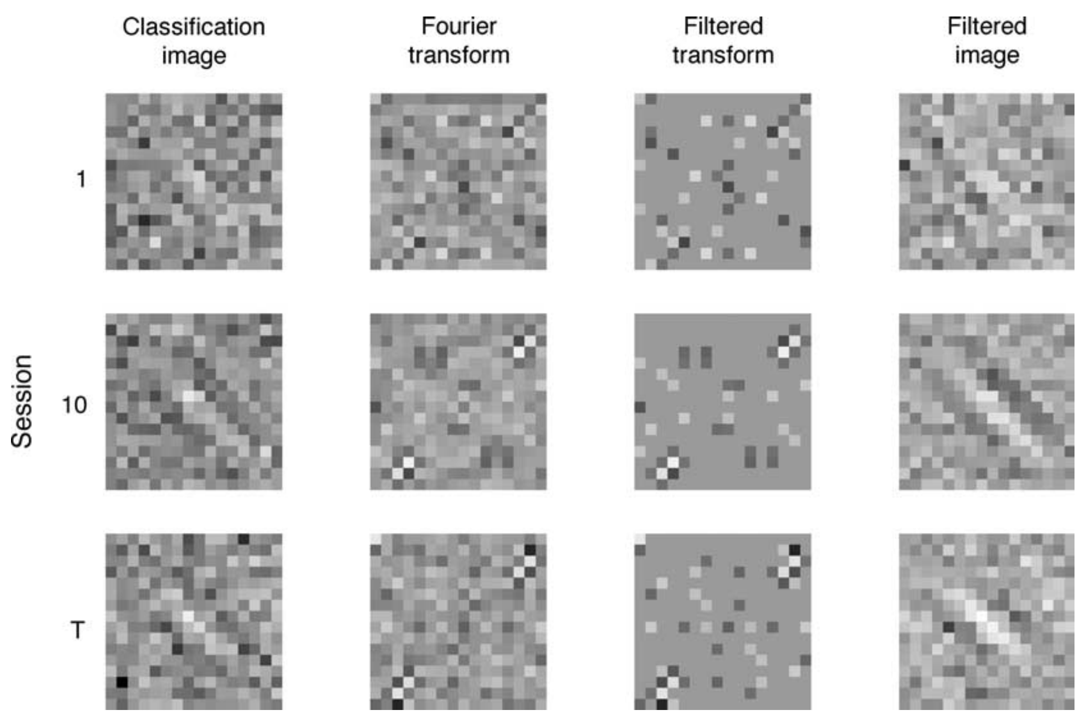

Figure 2. Classification images from a representative subject (first column). For the purposes of analysis and display, all image data have been rotated such that the target is oriented to 45°. Numbers denote training sessions and "T" denotes the transfer session. Images were run through a Monte Carlo filter in Fourier space to enhance qualitative features (fourth column), though this processing was not employed in formal image analysis.

custom distributions and converted to a normalized z-score, giving insight into which aspects or areas of the image are most important to observers.

Results

Image correlations

Classification images were computed using a linear regression algorithm (Victor, 2005). Images for one subject are shown in Figure 2. The more common "mean difference" method (Ahumada, 2002) was considered for use, but in our testing this algorithm's results were overly noisy and failed to produce clearer images even with very large numbers of trials (see Supplementary Figure 1). The linear regression algorithm was found to be robust against the effects of signal insertion and to produce clearer images with increasing numbers of trials, making it the superior solution for the stimuli used in this particular experiment. The correlation between the resultant classification image and an image of the target signal was computed as the primary learning metric. In this way, a higher correlation can be interpreted as a "better" classification image.

Classification image correlation plots are shown in Figure 3. Subjects show a significant improvement in image correlation of approximately 7% over the course of training (p = 0.012, one-tailed t-test, Day 1 vs. Day 10). Furthermore, the transfer session shows a significant decrease in the strength of image correlation (training session 10 vs. transfer session, p = 0.031). Taken together, these results suggest that perceptual learning is specific to the orientation of training.

Behavioral performance

We also examined traditional measures of perceptual learning. On average, subjects showed only a 3% improvement in threshold over training (Figure 4), and this effect was not significant (p = 0.25, Day 1 vs. Day 10, one-tailed t-test; p = 0.25, Spearman Rank Correlation between session and threshold). However, as discussed below, there were a number of subjects who showed performance decrements across sessions. Additionally, the threshold level of the transfer session shows no significant relationships.

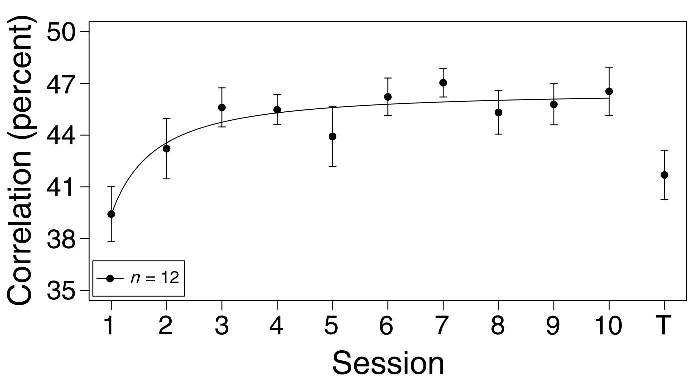

Figure 3. Correlation of the classification image with the target image. X-axis denotes training session number ("T" denotes transfer session, where stimulus was presented at a novel orientation orthogonal to the orientation of training). Error bars represent within-subject standard error.

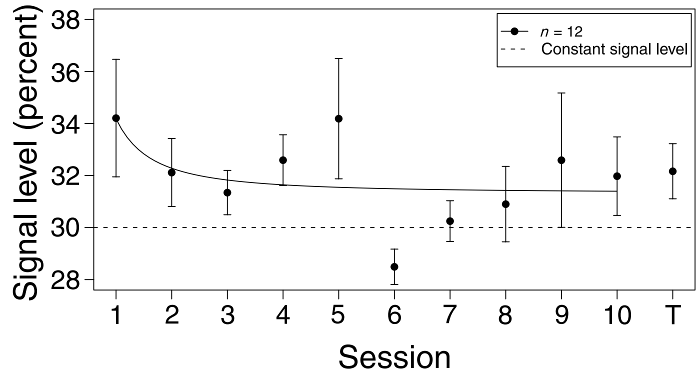

Figure 4. Average threshold. Dashed horizontal line denotes the signal percentage used in trials not controlled by the staircase.

Since we employed a mixed design in which half of the trials contained a stable signal level, we are also able to gauge improvements in accuracy. Subjects also showed a 3% improvement in accuracy on the constant signal level trials, which was significant (p = 0.035) and showed a significant correlation with session of training (r = 0.19, p = 0.036). Reaction times decreased significantly (p = 0.001) by an average of 160 ms during training.

Relationship between accuracy and classification image

We observed a strong negative correlation between the average threshold and the average session image correlation with the target image (r^2 = 0.51, p < 0.001). An

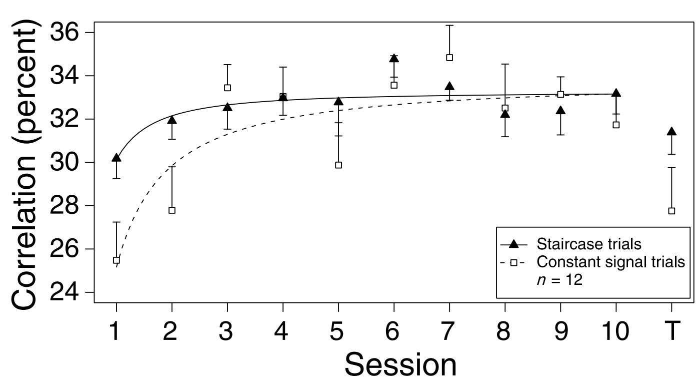

Figure 5. Correlations of classification images to the target image, produced separately from staircase-controlled trials and constant signal trials.

important question is whether the improvements in the classification images were simply related to the number of noise pixels in our images, which changes systematically with threshold. To control for this, we computed classification images separately for the constant signal and staircase-controlled stimulus sets (Figure 5). Both sets of images demonstrate clear learning effects between the beginning and end of training (p = 0.02 and p = 0.025, respectively), indicating that while there is a strong relationship between signal level and the strength of the classification image, the relationship is non-causal. Notably, examination of individual subject data verifies this non-casual relationship between threshold and the goodness of the classification images.

Orientation tuning functions

Just as the image correlation measure can be computed by correlating a classification image with the target signal image, it is possible to establish orientation tuning functions for each classification image by correlating it with target gratings at all possible orientations. Gaussians were then fit to these resultant tuning functions (see Figure 6). The height of the tuning function increased by approximately 9% on average (measured on the scale of image correlation), a significant increase (p = 0.003), and this improvement did not transfer to the untrained orientation in the transfer session (Day 10 vs. Transfer session, p = 0.008). Furthermore, the center of the tuning function shifted toward the target's orientation during training (p = 0.033), and this improvement appears to have transferred to the orthogonal orientation (p = 0.383). Although a few individual subjects demonstrated changes in tuning bandwidth, the effects were inconsistent and not significant (p = 0.219). In addition to the improvement in image correlation at the trained direction, there appears to be a strong inhibition of correlation surrounding the peak. This inhibition appears to be at least weakly present everywhere beyond 15 degrees of the trained orientation.

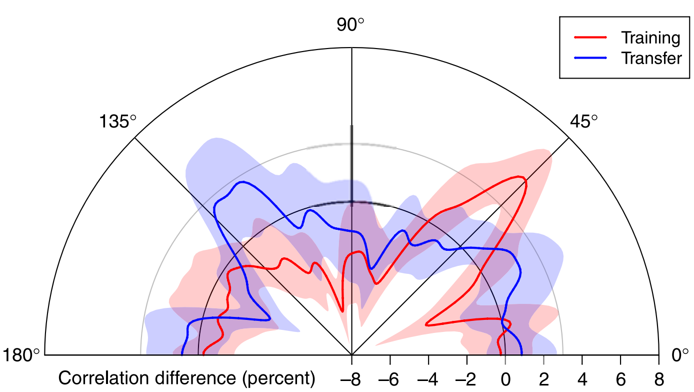

Figure 6. Orientation tuning functions for the final training session (red line) and the transfer session (blue line). See Methods section and Results section for further details. These functions are produced by subtracting the subject's baseline tuning function (Day 1) from the Day 10 and Transfer session functions. Owing to aliasing inherent in these low-resolution images, data have been smoothed with a Gaussian filter (σ = 3°). The gray line at 3% has been added to aid the visibility of the difference between the heights of the curves.

Individual differences

While the aggregate data show significant learning, we observed a high degree of between-subject variability in the learning effects. To clarify these individual differences, subjects were split into two groups based on how their thresholds changed over the course of training. Subjects whose thresholds had decreased by the end of training were classified into one group, deemed "learners," while those who did not were classified as "non-learners". Under these criteria, 6 subjects were classified as learners and 6 as non-learners.

Results of these two groups are summarized in Figure 7. As expected from the selection criteria, the learners also showed a 10% decrease in threshold over training (p = 0.001) and a significant correlation between session and threshold ($r = -0.39$, p = 0.002). On constant signal trials, learner accuracy improved by 7.5% (Day 1 vs. Day 10, p = 0.004) and there was a significant correlation between session and accuracy (r = 0.36, p = 0.004). Furthermore, reaction time decreased by 265 ms over training ($p < 0.001$). In addition, the learners showed a 13% improvement in image correlation over training (p = 0.004). The improvements in image correlation and accuracy were specific to the stimulus orientation (p = 0.015 and p = 0.04, respectively) and was close to significance for threshold (p = 0.08).

In contrast, the non-learner group shows no significant effects other than a 131-ms reduction in reaction time over training (p = 0.019). All other tests within the non-learner group are non-significant. They only showed a 1% improvement of image correlation over training (p = 0.705), a 6% increase of threshold (p = 0.144), and 1% reduction in accuracy over training (p = 0.646). Of note, these results show an important dissociation between different measures of performance, while most of the effects are non-significant, non-learners showed worse performance (i.e., higher thresholds and lower accuracy) but improved classification images and faster reaction times.

Average classification images for each group are shown in Figure 8. While in the first session the average images appear similar between groups, by Day 10 the image correlations were more pronounced for the learners compared to the non-learners. We also looked at changes in pixel intensity across the classification images for each group (Figure 9). Non-learners more consistently prioritized a central area of the stimulus, while learners based their decisions on a larger area of the stimulus as training progressed. From Days 1 to 10, pixel intensity increased

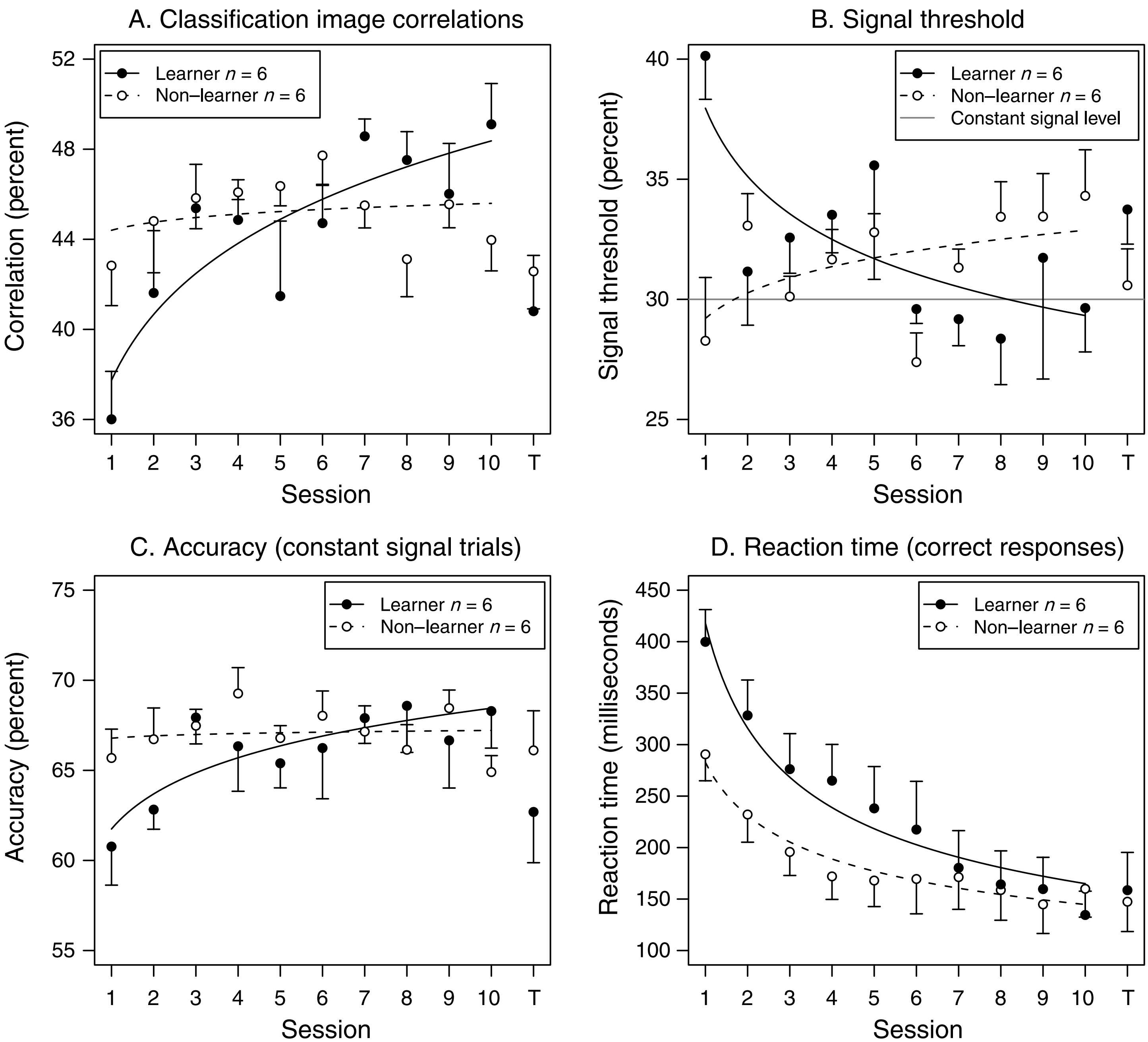

Figure 7. A comparison of learner and non-learner groups on the measures of (A) classification image strength, (B) threshold signal level, (C) response accuracy on constant signal trials, and (D) reaction time on correct responses.

significantly in areas surrounding the center of the stimulus for learners (see "Width" clusters 2 and 3 in Figure 9, p = 0.010 and 0.017, respectively), as well as more peripheral areas along the length of the stimulus ("Length" clusters 3 and 4, p = 0.005 and p = 0.021). In contrast, non-learners show no significant changes in pixel intensity across training sessions or pixel locations. These results show that an important aspect of learning was the increased reliance upon the more peripheral parts of the stimuli.

We also examined the extent to which subjects used the white vs. the black pixels in the image to discriminate the targets. To do this, we separately took image correlations for the bright areas versus dark areas of the target template. We found that all subjects in the learner group improved their "dark" correlations during training (mean improvement of 16.6%, p = 0.025), and 5 of 6 learners also improved in "bright" correlation (mean improvement of 8.2%, p = 0.011). In the non-learner group, 4 of 6 subjects improved in dark correlation (mean improvement of 7.3%, p = 0.131), 1 improved in bright correlation (mean decrement of −5.2%, p = 0.127), and no non-learner improved in both correlation types simultaneously (see Figure 10). These results suggest that one of the

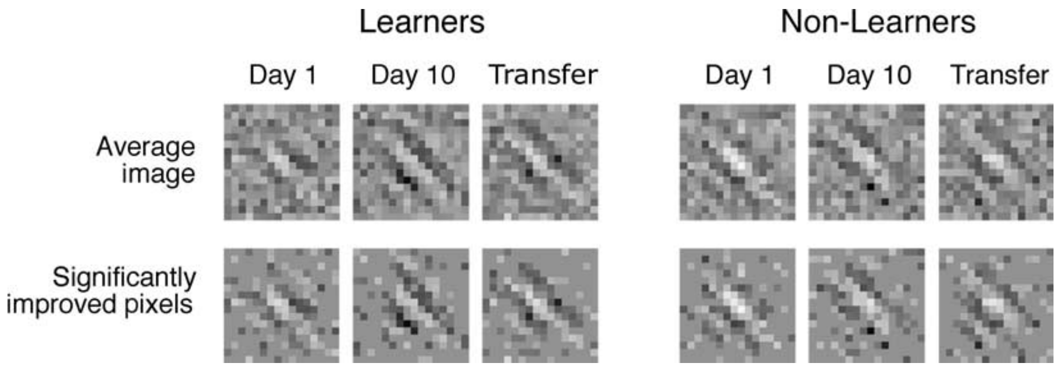

Figure 8. A comparison of image averages between learner and non-learner groups. The top row shows the average classification images for each group. The bottom row removes pixels that on average were not significantly greater than the mean of the image's distribution of possible values (see Methods section).

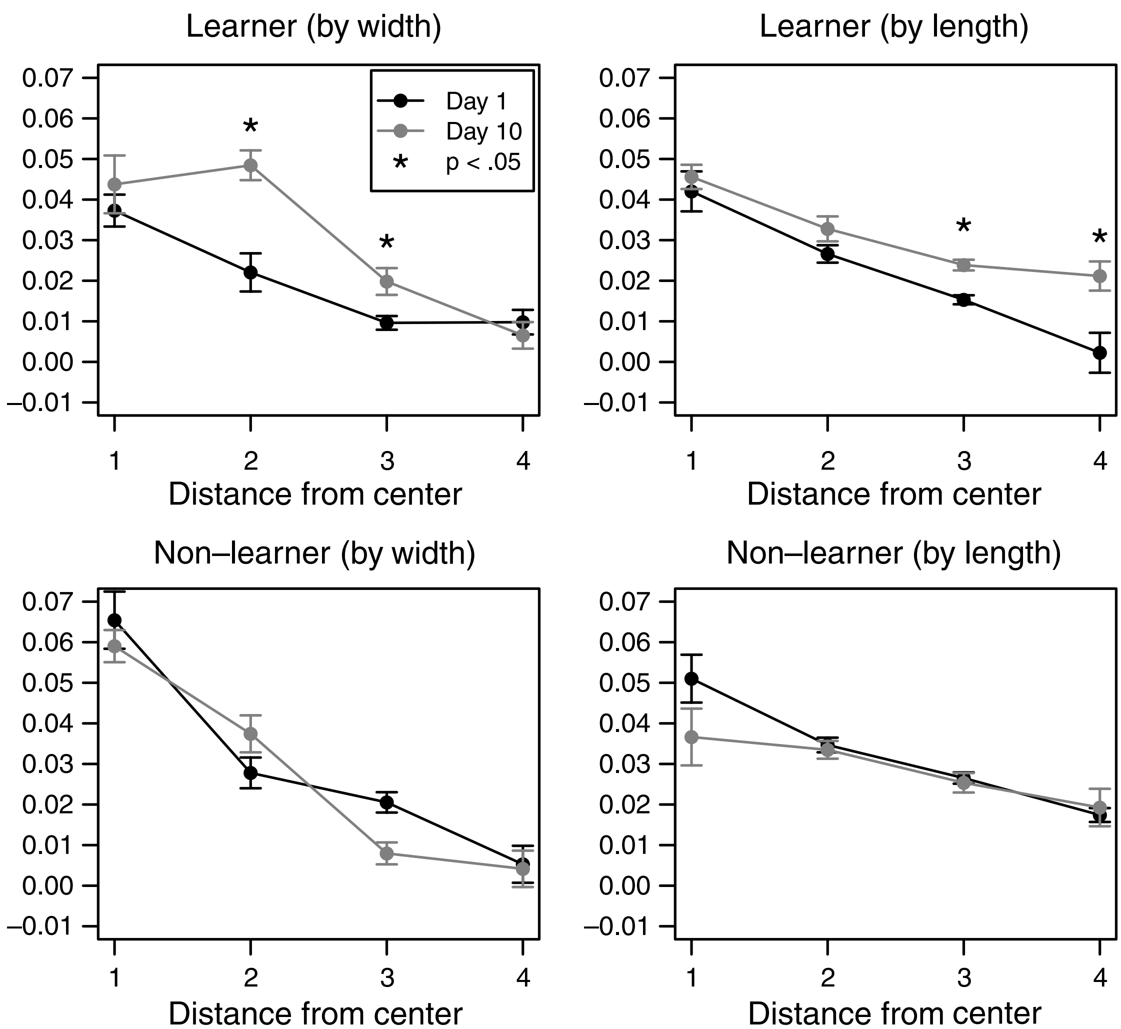

Figure 9. Changes in the length and width of classification images. Average pixel intensity is plotted for learners and non-learners on Training Days 1 and 10, proceeding from the center of the image (1) to the edge (4). Images are folded along the orientation of the grating (Width) or perpendicular to the orientation of the grating (Length) and intensities are averaged in bands that proceed along the grating's width or length, respectively.

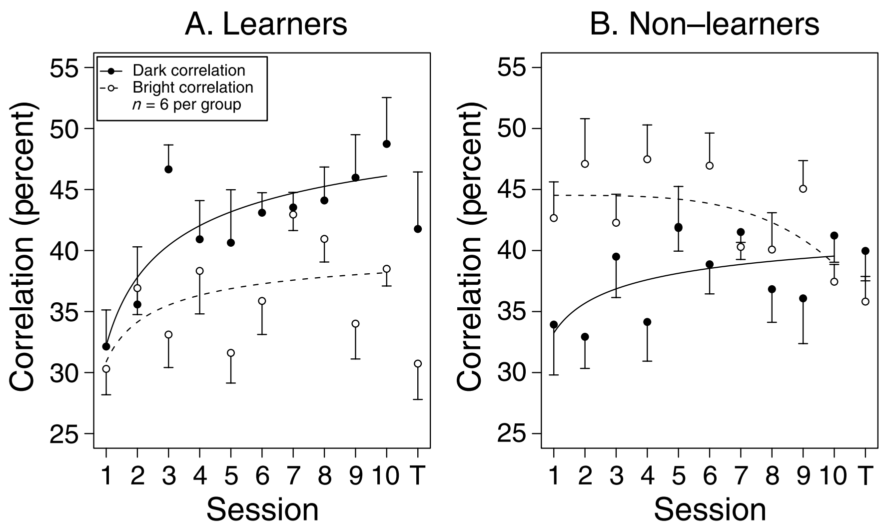

Figure 10. Changes in bright and dark correlations (comparing the classification image to bright and dark sections of the target template) for learners and non-learners.

reasons why the non-learners might not have learned is that they applied an inefficient strategy of concentrating on the dark bands of the image at the expense of the white bands (this can also be seen in Figure 9B where the non-learners get worse at the 1st and 3rd bands, which are white, and a little better at the 2nd band, which is dark). On the other hand, the learners were able to improve their performance on bands of both polarities as well as overall achieving a wider and longer classification image.

We observed that on average non-learners started off a little better on the task than the learners. Other studies have also found that initial performance on a task can be an important prediction of what subjects learn (Aberg & Herzog, 2009). To address this point, we compared baseline performance with metrics of performance change. These ratios reveal that baseline threshold and baseline accuracy are more predictive of overall learning for the non-learner group than the learner group (Supplementary Figure 2). Non-learners who had lower baseline thresholds tended to show less learning, while learners showed comparable levels of learning irrespective of baseline performance.

Discussion

The results of this study demonstrate that classification images for the discrimination of an orientation pattern in noise can be generated from a single session's data and successfully used to track perceptual learning. An advantage of classification image analysis is that the primary output of each session is an image, as opposed to a numerical representation of percent correct or noise threshold. This allows for a wide variety of image-based analyses that enable us to examine the effects of perceptual learning in greater detail than would otherwise be possible. Thus, we could go above and beyond confirming previous studies showing the orientation specificity of learning to identify individual differences of what subjects learned. For example, we observed that subjects in the learner group learned to use pixels across a wider area of the image, and, apposed to non-learners, showed improvements in both the white and dark parts of the image. These observations would not be evident through analysis of standard performance measures. Furthermore, we established orientation tuning curves for each classification image, allowing us to examine not just the intensity of observers' mental templates but their precision as well. These results show the advantages that classification images have in understanding what subjects learn (or do not learn) in the course of perceptual learning.

An important observation of our study is that measurements of accuracy, signal level, reaction time, and image correlation do not always track perfectly with one another, and evaluating multiple measures of performance can be diagnostic of different aspects of learning and of individual differences in strategies and what is learned. For instance, while learners and non-learners show different signal level changes across sessions, reaction times for the two groups follow very similar trends. This benefit of reaction time may reflect aspects of task-learning such as learning the timing of the stimuli, how to press keys more quickly, and generally becoming more comfortable with the experimental setting. Furthermore, while in the learners we found a general improvement in all metrics of performance, in the non-learners and in individual subjects we noticed several examples where thresholds and accuracy decreased when there were improvements in the classification images. These types of discrepancies may indicate that the non-learners were in fact learning, but that this learning involved trade-offs that privileged some aspects of the stimuli while neglecting others, such as the observed shift in focus from the bright to the dark regions of the image. Our results thus demonstrate that different learning metrics can reveal (or cannot reveal) different aspects of what subjects learn.

In order to generate sufficient data to resolve a classification image in a single session, stimuli were constructed with low dimensionality in mind. Stimulus resolution was reduced to 16×16 pixels, minimizing the number of data points comprising an individual stimulus. M-sequences were used to ensure a low level of cross-correlation in the set of noise masks being used. Pixel substitution methods and an aliased square-wave target were employed to eliminate shades of gray from all stimuli, ensuring that each pixel of a stimulus would contribute maximally to the final classification image. Using these designer stimuli, the linear regression algorithm was able to resolve a stable image in as few as 500 trials (Supplementary Figure 1). In this manner, a classification image could easily be produced from 1 h's worth of data, allowing phenomena such as perceptual learning to be examined with this technique.

Still, future further optimizations could be done to achieve an even more efficient paradigm. For example, our attempts to guarantee a low-dimensional stimulus set may have been counterproductive to a certain extent. A set of 256 unique noise masks was generated from the m-sequence. This set was then duplicated to produce a total of 1,024 masks. Half of these were then polarity-inverted to minimize mean luminance perturbations in the noise mask set. Further simulations suggest that the linear regression algorithm does not perform well when a large number of stimuli covary perfectly (as any given mask would with its own duplicates and polarity-inverted counterparts). Therefore, it may be possible to produce even better classification images if this particular aspect of our paradigm is discarded in favor of including a more diverse mask set.

We find that subjects' classification images improve significantly over the course of training. Observers demonstrate specificity for the trained stimulus consistent with other measures of perceptual learning. The significant decrease of threshold during training for the learner subjects suggests either that the relevant features of the template are strengthened or that internal noise is reduced (Dosher & Lu, 1998). A further addition to the paradigm that we introduce here would be to use a double pass method (Gold, Bennett et al., 1999; Li et al., 2004), where in essence each image with noise mask is presented twice, to further characterize aspects of perceptual learning. While there were some repetitions of stimuli in our experiment, these were typically not presented at the same signal level and this prevented us from reliably measuring internal noise of the subjects in our study.

A previous study using classification images to study perceptual learning (Dupuis-Roy & Gosselin, 2007) found that perceptual learning could occur and classification images could be generated from trials in which no explicit stimulus was presented (i.e., zero-signal trials). While this is a method in which one could generate classification images without any potential interference from the presence of an actual stimulus, we had difficulty getting subjects to perform well for zero-signal stimulus presentations (unpublished pilot data). We found that subjects reported a high degree of frustration with the task and resorted to random guesses. Thus, we found it necessary to insert a percentage of the target signal into the stimuli. While there is a strong correlation between signal percentage and final image strength, our split-session analyses suggest that this is not a dependent relationship.

Other studies have also observed individual differences in perceptual learning. For example, a recent study found that initial performance on a task can be an important predictor of the degree of specificity of the subsequent learning (Aberg & Herzog, 2009). Classification images give us another window into individual differences beyond these numerical analyses. Examining changes in pixel intensity in the classification images allows us to get an idea of where subjects were focusing their attention as they performed the task (Figure 9). Non-learners consistently prioritized a narrower portion of the stimuli during all sessions, while learners incorporated a larger area of the stimulus as training continued. Numerical data indicate that non-learners had lower baseline thresholds and higher baseline performance and produced seemingly "better" baseline images. However, it is clear from the pixel intensity data that this is a byproduct of a lazier perceptual strategy that produced stronger results in the short term but failed to produce meaningful learning in the long term.

Consistent with this finding, it is also clear from the data that learners and non-learners employed different strategies in regard to the bright and dark sections of the target stimulus. Learners gradually emphasize dark portions of the stimulus in addition to the bright, whereas non-learners seem to trade one for the other. Whether this strategic shortfall is the result of poor subject motivation or differences in the visual system is unclear, but such results point toward the richness of the classification image technique.

Conclusions

We have demonstrated an efficient variation on classification image analysis that is appropriate for the study of perceptual learning. The images produced from these methods are sufficiently stable to demonstrate the effects of perceptual learning on the detection of noisy oriented gratings. Observers performing this task demonstrate a clear improvement in the strength of the perceptual template during training. This improvement shows specificity consistent with other perceptual learning paradigms. The nature of the image-based output metric allows for a more detailed data analysis than is possible with standard performance metrics and can better reveal individual subject differences in perceptual learning. However, while classification images confer advantages in understanding individual subject differences in learning, image-based analysis of learning also poses analytic complexities and there is a need to further develop appropriate metrics by which to more closely examine and quantify subtleties of what individuals learn.

Acknowledgments

We would like to thank Tiffany Sanchez, Daniel Khafi, Ted Yau, Richard Wald, and Elizabeth Tiglao for their help collecting data in this experiment. JD was supported by CELEST, a National Science Foundation Science of Learning Center [SBE-0354378].

Commercial relationships: none.
Corresponding author: Aaron R. Seitz.
Email: aseitz@ucr.edu.
Address: Department of Psychology, University of California, Riverside, 900 University Ave., Riverside, CA 92521, USA.

References

Aberg, K. C., & Herzog, M. H. (2009). Interleaving bisection stimuli—randomly or in sequence—does not disrupt perceptual learning, it just makes it more difficult. *Vision Research, 49,* 2591–2598.

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences, 8,* 457–464.

Ahumada, A. J. (2002). Classification image weights and internal noise level estimation. *Journal of Vision, 2*(1):8, 121–131, http://www.journalofvision.org/content/2/1/8, doi:10.1167/2.1.8. [PubMed] [Article]

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

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.

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

Dupuis-Roy, N., & Gosselin, F. (2007). Perceptual learning without signal. *Vision Research, 47,* 349–356.

Fahle, M. (2004). Perceptual learning: A case for early selection. *Journal of Vision, 4*(10):4, 879–890, http://www.journalofvision.org/content/4/10/4, doi:10.1167/4.10.4. [PubMed] [Article]

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

Gibson, E. J. (1963). Perceptual learning. *Annual Reviews in Psychology, 14,* 29–56.

Gold, J., Bennett, P. J., et al. (1999). Signal but not noise changes with perceptual learning. *Nature, 402,* 176–178.

Keane, B. P., Lu, H., et al. (2007). Classification images reveal spatiotemporal contour interpolation. *Vision Research, 47,* 3460–3475.

Law, C. T., & Gold, J. I. (2008). Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nature Neuroscience, 11,* 505–513.

Li, R. W., Levi, D. M., et al. (2004). Perceptual learning improves efficiency by re-tuning the decision 'template' for position discrimination. *Nature Neuroscience, 7,* 178–183.

Lu, H. J., & Liu, Z. L. (2006). Computing dynamic classification images from correlation maps. *Journal of Vision, 6*(4):12, 475–483, http://www.journalofvision.org/content/6/4/12, doi:10.1167/6.4.12. [PubMed] [Article]

Mareschal, I., Dakin, S. C., et al. (2006). Dynamic properties of orientation discrimination assessed by using classification images. *Proceedings of the National Academy of Sciences of the United States of America, 103,* 5131–5136.

Mollon, J. D., & Danilova, M. V. (1996). Three remarks on perceptual learning. *Spatial Vision, 10,* 51–58.

Nishina, S., Seitz, A. R., et al. (2007). Effect of spatial distance to the task stimulus on task-irrelevant perceptual learning of static Gabors. *Journal of Vision, 7*(13):2, 1–10, http://www.journalofvision.org/content/7/13/2, doi:10.1167/7.13.2. [PubMed] [Article]

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Reid, R. C., Victor, J. D., et al. (1997). The use of m-sequences in the analysis of visual neurons: Linear receptive field properties. *Visual Neuroscience, 14,* 1015–1027.

Schoups, A., Vogels, R., et al. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature, 412,* 549–553.

Seitz, A. R., & Dinse, H. R. (2007). A common framework for perceptual learning. *Current Opinion in Neurobiology, 17,* 148–153.

Seitz, A. R., Kim, D., et al. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron, 61,* 700–707.

Sekuler, A. B., Gaspar, C. M., et al. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology, 14,* 391–396.

Shimozaki, S. S., Chen, K. Y., et al. (2007). The temporal dynamics of selective attention of the visual periphery as measured by classification images. *Journal of Vision, 7*(12):10, http://www.journalofvision.org/content/7/12/10, doi:10.1167/7.12.10. [PubMed] [Article]

Victor, J. D. (2005). Analyzing receptive fields, classification images and functional images: Challenges with opportunities for synergy. *Nature Neuroscience, 8,* 1651–1656.

Xiao, L. Q., Zhang, J. Y., et al. (2008). Complete transfer of perceptual learning across retinal locations enabled by double training. *Current Biology, 18,* 1922–1926.