

Partial Position Transfer in Categorical Perceptual Learning

Alexander Gerganov (agerganov@cogs.nbu.bg)

Department of Cognitive Science and Psychology, New Bulgarian University
21 Montevideo St., 1618 Sofia, Bulgaria

Maurice Grinberg (mgrinberg@nbu.bg)

Department of Cognitive Science and Psychology, New Bulgarian University
21 Montevideo St., 1618 Sofia, Bulgaria

Robert L. Goldstone (rgoldsto@indiana.edu)

Department of Psychological and Brain Sciences, Indiana University
1101 East Tenth Street, Bloomington, IN 47405 USA

Abstract

Two experiments are reported. The first shows incomplete transfer of explicit categorical learning at a distance of 4.5 degrees of visual angle and the second is a control experiment with a non-learning task. The results suggest that some early visual plasticity takes place even in a simple, explicit categorical learning task. We claim that perceptual learning is a much more common phenomenon than believed before and that it plays an important role in everyday tasks including higher-level learning.

Keywords: perceptual learning; categorization; translation invariance

Introduction

Perceptual learning is often defined as performance improvement in simple discrimination tasks that involve neural structures dedicated to low-level perceptual processing. For the visual modality such tasks are Vernier discrimination, line orientation discrimination, and simple pattern recognition. Due to the nature of these tasks, participants rarely report any explicit strategy change over the course of improvement, thus perceptual learning is considered to be implicit. The assumption for implicitness of perceptual learning, however, is more or less intuitive and presumed on the basis of the specific nature of the learning tasks that are used. In real life pattern recognition, line discrimination and other similar visual tasks are usually small parts of more complex cognitive tasks like categorization for example. Categorization is a versatile task that spans perceptual tasks such as identification and segmentation as well as higher-level reasoning and cognitive strategies. As such, categorization is a good task to investigate what is the role of perceptual learning in a complex task in which explicitness, complexity and other parameters can be varied so that we can better understand the interaction between perceptual and conceptual processes.

A point of disagreement among researchers is the exact locus of perceptual learning. One view is that perceptual learning results from long-lasting changes that occur at a relatively early stage in information processing from sensory registration to decision making (e.g. Fahle, 1997;

Karni & Sagi, 1991). Another position is that only the projections from the low perceptual level to higher cognitive levels change in the course of learning, with the early visual representations remaining constant (Dosher & Lu, 1999).

One reason for this debate is that different tasks lead to results that contradict each other. Another reason, however, could be that the notions of “low-level perception” or “higher-level decision making” are seldom grounded in real brain structures. For vision, where the sensory pathways are relatively well understood, any theory or model about low-level perception could attempt to relate directly to concrete structures like V1, V4, inferotemporal cortex, etc. Experimental evidence for the functions of these structures comes from neurophysiological as well as behavioral studies. The translation invariance paradigm supplies a behavioral approach to testing the particular level of neural structures involved in a perceptual learning task (Notman, Sowden & Özgen, 2005; Sowden, Rose & Davies, 2002).

Translation invariance— theoretical rationale

The translation invariance paradigm is based on our contemporary knowledge about the visual pathways and the cellular structures involved in early visual information processing. The receptive fields of neurons in the visual pathways vary in size with the tendency to become larger in the higher visual areas (see Figure 1).

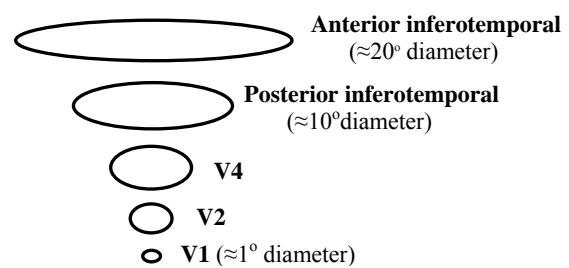


Figure 1: Receptive field size of neurons found in different parts of the visual pathways (not drawn to scale)

Neurons in V1 have receptive fields of only 1-2 degrees of visual angle, V4 is characterized by cells with larger receptive fields (4-5 degrees) and neurons in the inferotemporal cortex can have receptive fields of 10-20 degrees.

Translation invariance tests if successful learning, achieved in a particular location in the visual field, continues to be shown when the same stimuli are presented at another location (see figure 2). In this paradigm, participants are asked to look at a fixation point in the center of a screen while stimuli are presented for 100ms in the parafoveal vision. Given that 100ms is not sufficient time to initiate a saccade (Saslow 1967), only a small, specific region of the retina is stimulated. When performance meets a predefined criterion, the same stimuli are presented at a new location in the visual field. Theoretically there are three possible outcomes. The performance could remain the same as before the shift, indicating that the learning is invariant to translation over the retina. The performance could drop to its initial, pre-training level, indicating that the learning is position specific. Finally, there could be partial transfer of the learning if performance declines somewhat at the new location, but not to its initial level.

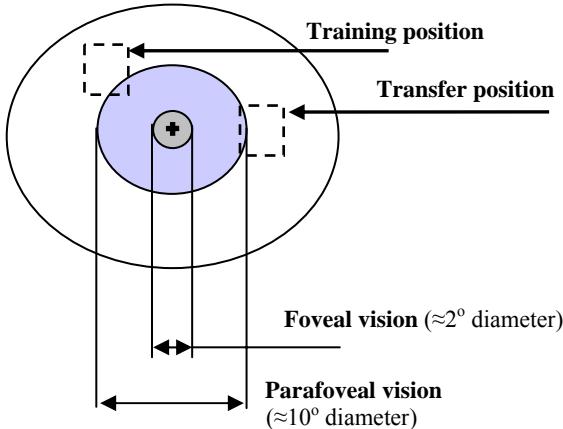


Figure 2: Translation invariance paradigm – the visual field and presentation of stimuli at particular locations.

Given the systematic relation between receptive field size of neurons and their cortical location, these three outcomes provide evidence for the location of visual structures that take part in the learning process. In the first scenario, where full transfer is observed, the receptive fields of the involved structures should cover both the training and the transfer location, indicating neural change at the level of posterior Inferotemporal cortex or later. If there is no transfer at all, probably the learning occurred at a neural level comprising neurons with smaller receptive fields than the distance between the training and transfer locations (see figure 3), implicating an earlier site of plasticity.

There are many studies that have found no transfer or partial transfer of learning for simple tasks like Vernier discrimination (e.g. Fahle, Edelman & Poggio, 1995), line orientation discrimination (Shiu & Pashler, 1992) pattern

discrimination (Dill & Fahle, 1997) and other. This is interpreted as evidence for a very early locus (V1-V4) of perceptual learning (but see Mollon and Danilova, 1996 for criticisms of this assumption.

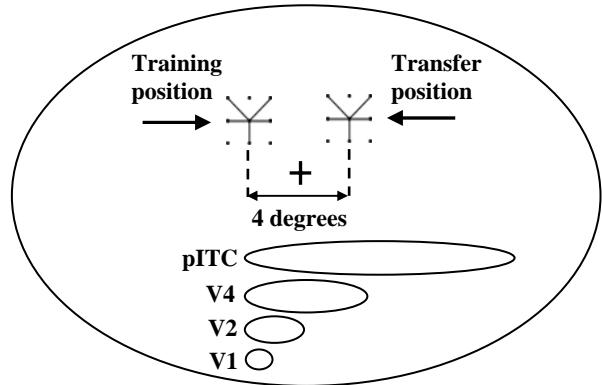


Figure 3: Translation invariance vs. position specificity of learning. Distance between the training and transfer positions related to the size of receptive fields of different types of cells.

Studies show that the degree of translation invariance depends on several factors: the symmetry between the training and transfer locations – translation invariance was found for symmetric transfer positions around the fovea (both on the X and Y axis) but not for asymmetric positions; the complexity of the task – the more complex and demanding the learning task, the more position-specific the learning is; the distribution of attention resources; and the previous exposure of the participants to parafoveal tasks (see Dill, 2002 for a review).

Experiments

In order to systematically test the perceptual learning component in categorical learning, we started with the simplest possible experiment, which acts as a baseline for further studies. The task was relatively simple (completing the experiment took between several minutes and half an hour); learning was fully explicit (manipulated by the instructions and tested in a post-study questionnaire). If explicit categorical learning involves only areas above V4, learning should be fully translation invariant.

Experiment 1

Method In this experiment we tested transfer performance as a function of the distance from the training location. Two transfer locations (at distances of 2 and 4.5 degrees of visual angle) were compared to the training location (control condition). We measured performance as the percent of correct responses and the response time for giving an answer (within-subject dependent variables). The order of

presentation of the control and two training conditions was fully counterbalanced.

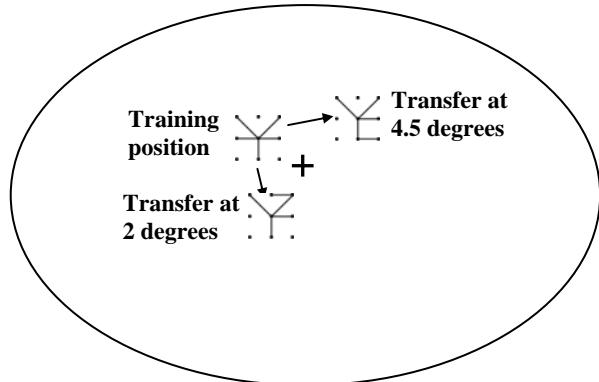


Figure 4: Training and testing locations.

The training position was located 2.5 degrees to the top and to the left from the fixation cross. The tested transfer locations were shifted by 2 and 4.5 degrees of visual angle from the training position. All three positions were at the same distance from the fixation cross and were not symmetric in any way (Figure 4).

Participants Participants in the study were undergraduate students at New Bulgarian University, who were recruited with print ads spread around the university and were paid for participation in the study. 24 participants finished successfully the training phase and were tested for translation invariance of learning. One outlier was excluded from the analysis because of a very low score (55%) in the control position. All other participants scored above 85% in the control position.

Apparatus For the translation invariance paradigm we used a Tobii 1750 remote eye tracker to ensure that participants were looking at the fixation cross in the center of the screen. The Tobii 1750 eye tracker provided a gaze sample every 20 ms and an E-prime 1.2 script checked in real time if the gaze was indeed at the fixation cross. Participants had to look continuously at the cross for 200ms for the stimulus presentation to be triggered. The eye tracker also checked that the viewing distance was between 58 and 62 cm. When these rules were satisfied, the cross disappeared from the screen and at the same time the stimulus appeared for 100ms at the appropriate location. Participants' responses and reaction times were collected using a Serial Response Box (SRBox), which guaranteed very low error in reaction time estimation (<1ms). Stimuli were presented on a 17" TFT screen (with a matrix response time of 25 ms) built in the Tobii 1750 eye tracker. We used the native screen resolution - 1280x1024.

Stimuli The stimuli were constructed in a fashion similar to Pevtzow and Goldstone's (1994) stimuli – they were composed of black line-segments connecting dots on a 3 by

3 grid (see Figure 5). The line segments were 2 pixels wide (approximately 0.05 degrees of visual angle) and the points were with diameter 3 pixels. The overall stimulus was 80 pixels high and 80 pixels wide (corresponding to approximately 2 degrees of visual angle). There were four basic stimuli which were distorted by adding a random line to the stem. The random lines were added to all possible places except if the added line resulted in any closed parts or detached lines. Categories were defined by a pair of overlapping lines shared by category members (see Figure 5). Participants were given at random one of two different categorization rules in order to avoid possible effects that were specific for the selected characteristic element. Participants in the experiment always saw one object at a time, composed of 5 lines (a four-line basic object plus a randomly added line). For every participant there were two categories and the same categorization rule remained throughout the whole experiment. The stimuli were presented on a white screen.

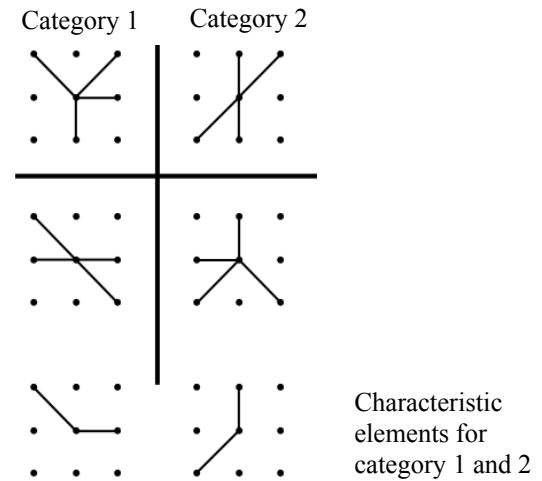


Figure 5: Stimuli – four basic stimuli and the characteristic elements

Procedure Participants were seated in front of the eye tracker (at 60 cm viewing distance from the screen) and were told that they would take part in an experiment studying cognitive processes during learning. They were given instructions about their task (categorization) and were shown four sample stimuli with the characteristic elements marked (the marked elements were different from the actual ones used throughout the study). The participants were told that during the experiment they would have to find by themselves which correct characteristic elements determined the categorization, but that the elements would be the same throughout the whole experiment. After reading the instructions, participants were briefly familiarized with the stimuli and the way of presentation – participants were asked to look at the fixation cross in the center of the screen and when their gaze was at the cross, a stimulus was presented for 100ms at the training position and then

disappeared (there was no mask after it). During the familiarization, participants did not try to categorize the objects. They observed them passively. After the familiarization, participants were told they would have to categorize each object into either Category “1” or “2” by pressing the labeled buttons on the SRBox. After pressing the desired button, they received immediate feedback as to whether their choice had been correct or wrong and their cumulative percent correct.

Trials were organized in blocks of 20. Participants had to score at least 90% correct answers during the 20 trials in order to continue to the next phase of the experiment. The percent of correct answers was computed after each trial and was displayed as feedback. After 20 consecutive trials, the participants were presented an additional message with their overall results for the last 20 trials. This message marked the end of a block and the beginning of the next one (the percent of correct answers was computed for each block).

When participants reached above 90% correct answers for a block of 20 trials, they were congratulated and were told they had passed the first part of the experiment and that they would continue the same task with the same rules and objects, but the objects were to be presented at different places on the screen. After this instruction there were three blocks of 20 trials each, every block was presented at one of the three locations (training position, 2 degrees transfer and 4.5 degrees transfer). The order of presentation was fully counterbalanced across subjects – for 3 positions there were 6 possible combinations and every participant was randomly assigned to one of the 6 combinations.

At the end of the experiment, the participants were asked about their strategies and if they said they were looking for a diagnostic part, they were asked to draw it.

Results

The post-experiment questionnaire showed that all but one of the participants were able to explicitly define their categorization strategy - looking for the presence of a diagnostic part in the objects and draw the characteristic element for which they were looking. Thus, the learning was indeed explicit.

The mean percent of correct responses in all conditions was above 90% (see Table 1 for details). Still, a Repeated Measures ANOVA showed there was a significant difference between conditions: $F(2, 46) = 7.813; p < 0.001; \eta_p^2 = 0.254$;

A Pairwise comparison of the means (with Bonferroni adjustment) showed that there was no difference between the control position and the 2 degrees transfer position ($p = 0.647$). There was difference between the control and 4.5 degrees transfer positions ($p < 0.05$) and between the 2-degrees-transfer position and the 4.5-degrees-transfer position ($p < 0.05$).

Difference between the conditions was also found in the reaction times. Repeated Measures ANOVA showed slower reaction time for the 4.5-transfer-position: $F(2, 46) = 3.476; p < 0.05; \eta_p^2 = 0.249$.

Table 1: Mean percent of correct responses per condition.

Position	Mean, %	Std. Error
Control	97.1	0.8
Transfer 2 degrees	95.6	1.3
Transfer 4.5 degrees	91.9	1.7

A Pairwise comparison (with Bonferroni adjustment) between the means shows that only the difference between the 2-degrees and 4.5-degrees conditions was significant: $p < 0.05$; see Table 2 for mean RTs.

Table 2: Mean reaction time per condition.

Position	Mean, ms	Std. Error
Control	830	54
Transfer 2 degrees	791	54
Transfer 4.5 degrees	919	66

d' was also computed as a more reliable, criterion-independent measure than the percent correct responses – the general pattern of results was the same for d' .

Discussion of Experiment 1

Results from Experiment 1 showed partial transfer of learning for the 4.5-degrees condition - there was a significant difference (in terms of both correct responses and response time) between the control position and the further transfer position (4.5 degrees shift). This could be interpreted as evidence that lower-level visual structures (V2-V4) play a role in the changes that occur during learning even in this simple category learning scenario.

Another explanation, however, could be connected to selective spatial attention (Dill, 2002). One natural way to control for attention would be to randomize the presentation of objects during the testing phase on a trial-by-trial basis. This, however, would have considerably altered the task during the testing phase (since there was only one location during the learning phase) and affect the overall performance for all locations (as pretests hinted). For similar reasons O’Toole and Kersten (1992) trained participants with two patterns at two different locations and indeed found that attention could explain positional specificity of learning in their case. Since such training would have made the task much more complex, we chose a different approach. A second experiment was performed to determine what the role of selective spatial attention is in a similar, but non-learning task.

Experiment 2

The second experiment was a control one, which aimed to rule out selective spatial attention as possible explanation of the results from Experiment 1. The experiment was identical to the previous one with only the instruction changed - participants were shown a characteristic element and were given the simple task to press Button 1 if the element was present in an object and Button 2 if it was not. This task imitated the strategy that all of the participants in Experiment 1 had reported but did not involve any learning. Different characteristic elements were given to different subjects at random. Everything else remained the same as in Experiment 1. If spatial attention was the explanation of the results from Experiment 1, then the same pattern should be observed in Experiment 2.

Participants 22 participants were recruited from the same pool as in Experiment 1.

Results

Participants achieved 90% correct answers very quickly (usually in the first block). Thus, this task involved no learning as intended. There was no significant difference between the control and transfer positions (see Table 3).

Table 3: Mean percent of correct responses per condition.

Position	Mean, %	Std. Error
Control	96.8	1.07
Transfer 2 degrees	98	0.78
Transfer 4.5 degrees	95	1.14

The overall model (with Greenhouse-Geiser correction) showed a trend: $F(2, 42) = 2.558; p=0.109, \eta_p^2 = 0.109$ but only because of the difference between the two transfer positions (the control being in the middle). Indeed, the Pairwise comparison (with Bonferroni adjustment) showed no difference between either of the transfer positions and the control one: $p = 0.512$ for the difference between the control and the 2-degrees-transfer position and $p = 0.741$ for the key comparison between the control and 4.5 degrees transfer, which was significant in Experiment 1. The reaction times, however, showed a similar pattern to the results in experiment 1 (see Table 4).

Table 4: Mean response time per condition.

Position	Mean, ms	Std. Error
Control	644	32
Transfer 2 degrees	643	28
Transfer 4.5 degrees	744	35

Repeated Measures ANOVA showed slower reaction time for the 4.5-transfer-position: $F(2, 42) = 15.248; p < 0.001, \eta_p^2 = 0.42$

Discussion of Experiment 2

The results from Experiment 2 showed no statistically significant difference for the percent of correct responses (unlike Experiment 1) but significant delay in the response time for the 4.5 degrees transfer (as in Experiment 1).

Correct responses, %

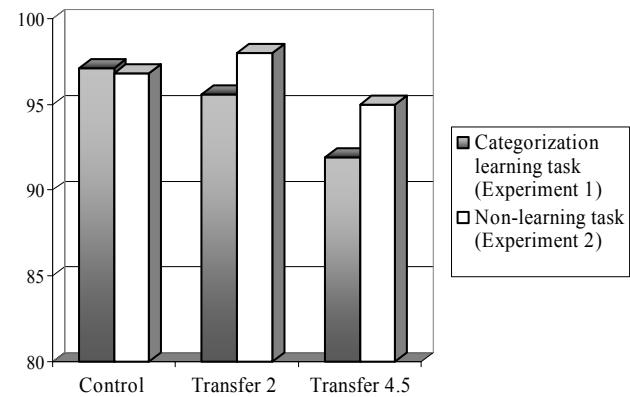


Figure 5: Percent correct answer for the control and two transfer positions, for experiment 1 and experiment 2

It is worth mentioning that the overall response times were much faster in the second experiment (see Figure 6), which is consistent with the accuracy results in indicating that the task was easier. Although the task imitated the strategy reported by participants in Experiment 1 (i.e. when a characteristic part is present press Button 1, when it is not – press Button 2) the faster response times show that the categorization task involved some additional cognitive processing.

Reaction time, ms

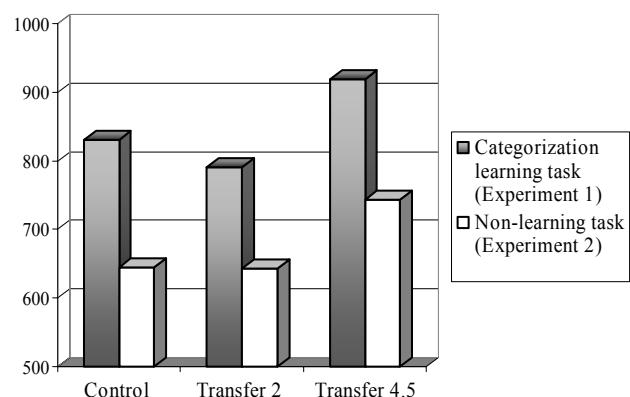


Figure 6: Reaction times (in ms) for the control and two transfer positions, for experiment 1 and experiment 2

General Discussion

The first experiment showed some evidence for involvement of low-level structures (as low as V2-V4, but not V1) in the category learning which supports and builds on previous findings (Sowden & Notman, 2005). Selective spatial attention could be another explanation for these results. The second experiment, however, showed that there was no effect of attention on position performance (in the percent correct responses) for a non-learning but otherwise identical situation. The delay in response time for the 4.5 transfer was similar to the observed in Experiment 1 and this effect could indeed be caused by selective spatial attention. The percentage of correct responses, however, did not match the pattern from Experiment 1, indicating that in addition to attention, spatially restricted perceptual learning is occurring. One possible interpretation of these results is that the learning process actually occurs simultaneously at different perceptual stages. Indeed previous research (Pevtzow & Goldstone, 1994) has shown the formation of perceptual detectors in categorical learning. Interestingly, neurophysiological studies have shown the formation of similar diagnostic shape detectors in the inferotemporal cortex of monkeys (Baker, Behrmann & Olson, 2002). Given that the current category learning task seems to involve structures even lower (i.e. earlier in the information processing stream) than IT, this supports the idea that several low-level perceptual layers are altered in parallel during perceptual learning.

Contemporary perceptual learning models tend to assume that learning happens at one particular stage of processing. A recently suggested computational model, however, shows how learning at different perceptual stages can occur simultaneously (Gerganov, Grinberg, Quinn & Goldstone, 2007).

In conclusion, Experiment 1 explored the question of whether category learning involves strategic, high-level changes to the weighting given to stimulus components, or lower-level changes to how stimuli to be categorized are perceived. If one grants the logic of translation invariance (e.g. lower-level perceptual processes show more position sensitivity because early visual pathways are served by detectors with narrow receptive fields), then our results suggest that category learning does alter relatively low-level perceptual processes.

Acknowledgments

This research was funded by U.S. Department of Education, Institute of Education Sciences Grant R305H050116 and National Science Foundation Grant 0527920.

References

Baker, C. I., Behrmann, M., & Olson, C. R., 2002. Impact of learning on representations of parts and wholes in monkey inferotemporal cortex. *Nature Neuroscience*, 5, 1210-1216.

Dill, M. (2002). Specificity versus Invariance of Perceptual Learning: The Example of Position. In *Perceptual Learning* edited by Manfred Fahle, Tomaso Poggio

Dill, M., & Fahle M. (1997) Limited Translation Invariance of Human Visual Pattern Recognition. *Perception and Psychophysics*, 60, 65-81.

Dosher, B. A., & Lu, Z.-L. (1999). Mechanisms of perceptual learning. *Vision Research*, 39, 3197-3221.

Fahle, M. (1997). Specificity of learning curvature, orientation, and vernier discriminations. *Vision Research*, 37, 1885-1895.

Fahle, M., Edelman, S., & Poggio, T. (1995). Fast perceptual learning in hyperacuity. *Vision Research*, 35, 3003-3013

Gerganov, A., Grinberg, M., Quinn, P. C., & Goldstone, R. L. (2007). Simulating conceptually-guided perceptual learning. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the Twenty-ninth Annual Conference of the Cognitive Science Society* (pp. 287-292). Austin, TX: Cognitive Science Society.

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences, USA*, 88, 4966-4970.

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

Notman, L., Sowden, P. T., & Özgen, E. (2005). The nature of learned categorical perception effects: a psychophysical approach. *Cognition*, 95, B1-B14.

O'Toole, A. J., & Kersten, D. J., 1992, Learning to see random-dot stereograms. *Perception*, 21, 227-243

Pevtzow, R., & Goldstone, R. L. (1994). Categorization and the parsing of objects. *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society* (pp. 717-722). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Saslow, M. G. (1967) Latency for saccadic eye movement. *Journal of the Optical Society of America*, 57, 1030-1033.

Shiu, L., & Pashler, H. (1992). Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Perception and Psychophysics*, 52, 582-588

Sowden, P. T., Rose, D., & Davies, I.R.L. (2002). Perceptual learning of luminance contrast detection: specific for spatial frequency and retinal location but not orientation. *Vision Research*, 42, 1249-1258.

Sowden, P. T., & Notman, L. A. (2005). Categorical Perception: Categorisation Dependent Perceptual Learning [Abstract]. *Journal of Vision*, 5, 865,865a