PERCEPTUAL LEARNING OF COMPLEX PATTERNS

by

Zahra Hussain

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Department of Psychology, Neuroscience and Behaviour
McMaster University

Descriptive Note

Doctor of Philosophy (2009)
Psychology, Neuroscience and Behaviour
McMaster University
Hamilton, Ontario

TITLE: Perceptual learning of complex patterns

AUTHOR: Zahra Hussain, B. A. (Sophia College, Mumbai University)

SUPERVISORS: Professor Patrick J. Bennett & Professor Allison B. Sekuler

NUMBER OF PAGES: xiv, 130

Abstract

Perceptual learning of complex patterns

Zahra Hussain
Doctor of Philosophy
Department of Psychology, Neuroscience and Behaviour
McMaster University
2009

Practice improves sensory perception, a phenomenon known as perceptual learning. Perceptual learning is interesting because it reflects plasticity in the brain where none was imagined, and because of its enormous applied potential. In vision, learning of simple discriminations is well-described. Here, I study the learning of two complex visual tasks, texture- and face identification, using a ten-alternative forced-choice procedure. The data are clear: learning of complex patterns is much like learning of simple patterns in its specificity, stability and time-course. Therefore, learning obeys similar rules at several levels in visual processing. The characteristics of learning, in particular the specificity and stability of learning, affect inherent aspects of object recognition.

Acknowledgements

It was a joy to learn from Patrick Bennett and Allison Sekuler, my mentors. I am deeply grateful to them for their guidance and will always think of them with enormous respect and affection. Donna Waxman ran most of the subjects, which is hundreds of people, and I basically cannot thank her enough for that. Former and current lab members helped in many technical and non-technical ways. Jesse, in particular, sat by me the entire time. My parents, as always, gave me lots of support and love throughout (and they chipped in some bucks!). So many other lovely people...including my sisters and brother who were, and insist on being, present in everything. Extreme thanks to all.

Preface

All chapters in this dissertation, except the General Introduction and the Summary, have been published, or are being submitted for publication in peer-reviewed journals. My advisors Patrick Bennett and Allison Sekuler are co-authors on all published chapters. Throughout, my role consisted in the design and running of the experiments, data analysis, and preparation of the manuscripts in collaboration with my advisors.

Chapter 2: Inversion effects are learned: Specificity of perceptual learning for faces and textures, is in press in the journal Vision Research.

Chapter 3: Contrast reversal abolishes perceptual learning is published in the Journal of Vision, and is available at: http://journalofvision.org/9/4/20/

Chapter 4: Robust perceptual learning of faces in the absence of sleep, is published in Vision Research.

Chapter 5: How much practice is needed to produce perceptual learning? is in revision, Vision Research.

Chapter 6: Superior identification of familiar visual patterns a year after learning, is in the final stages of preparation for submission (journal to be determined).

Contents

Chapter 1

General Introduction

Practice improves sensory perception: where initially oblivious, we can learn to perceive subtle distinctions in auditory, tactile and visual information. The improvements on sensory tasks thus brought about are called perceptual learning, and occur through changes in the brain. Therefore, the most basic aspects of behaviour, such as seeing, are affected by the brain's capacity to be shaped by experience. Barlow recently wrote: "It is a mistake to consider perception and learning separately because what one learns is strongly constrained by what one perceives, and what one perceives depends on what one has experienced" (Barlow, 1990). Here, I consider the role of experience in the visual identification of complex visual objects.

In vision, there are numerous demonstrations of perceptual learning of basic discriminations involving simple stimuli (McKee and Westheimer, 1978; Fiorentini and Berardi, 1981; Ball and Sekuler, 1987; Schoups et al., 1995; Fahle et al., 1995; Karni and Sagi, 1993). Practice improves detection of offsets between lines (McKee and Westheimer, 1978), discrimination of Gabor orientation (Schoups et al., 1995), and discrimination of dot motion direction (Ball and Sekuler, 1982, 1987). A robust characteristic of perceptual learning with simple stimuli is stimulus-specificity: the effects of learning vanish when the stimuli are altered (Schoups et al., 1995; Ball and Sekuler, 1987), or shifted to a different location in the visual field (Fahle et al., 1995; Schoups et al., 1995; Karni and Sagi, 1991; Crist et al., 1997). This type of learning has been attributed to changes in the response properties of cells in primary visual areas, where elemental stimulus properties such as orientation are encoded, and where cells are retinotopically organized. In other words, stimulus-specific learning of simple visual discriminations is taken as evidence for plasticity of brain regions that represent basic attributes of the visual world (Karni and Bertini, 1997; Gilbert, 1994).

Perceptual learning has a distinct time-course that involves two components. Rapid within-session improvements occur early and are attributed to generalizable learning of task-demands; Gradual across-session improvements are measured between sessions and are thought to represent stimulus-specific processes, such as the structural modification of cell ensembles that encode the stimuli (Karni and Bertini, 1997). In certain cases, it has been shown that sleep must intervene between sessions to consolidate what has been learned (Karni et al., 1994; Stickgold et al., 2000). Learning is often stable across time, with the improvements intact several months to years later, which suggests lasting modifications to the neural networks recruited during the task (Ball and Sekuler, 1982; Fiorentini and Berardi, 1981; Karni and Sagi, 1993). Therefore the hallmarks of perceptual learning of most simple visual tasks are the specificity, time-course and stability of improvements.

Complex visual stimuli can also be learned, and there are several reasons to study this type of learning:

1. Objects with multiple features that vary about several dimensions are behaviourally relevant. We rarely interact with isolated dots, lines and bars. Therefore, it is worth knowing how complex patterns are learned. On the one hand, the type of learning found in reduced contexts might be unique to reduced conditions. On the other, the principles of learning might be similar across a range of stimulus and task complexities. Are the hallmarks of perceptual learning uniform across a variety of conditions?

2. Learning of complex objects, as with simple patterns, tells us about how those objects are represented. Improvements on simple discriminations are usually constrained to the trained values of a given dimension, implying unique encoding of those values. For example, practice-related benefits at detecting the leftward motion of dots, do not transfer to the rightward direction (Ball and Sekuler, 1987). This means that leftward and rightward motion are encoded by unique cells in that region of visual space. There are many other examples of specificity of learning of simple patterns, all of which have helped illuminate the resolution and range of low-level visual representations (Schoups et al., 1995; Karni and Sagi, 1991; Crist et al., 1997). Recognition of complex objects is robust to many of the variations that affect processing of simple stimuli. We know that a car is a car regardless of its size, orientation or viewpoint. This means that learning could transfer across changes to the size, orientation or viewpoint of a complex object,

but does it?

3. Certain classes of objects, faces for example, are thought to be treated preferentially by the brain. A massive literature on face processing has yielded many behavioural and neural indices of supposedly face-specific processing: the face-inversion effect (Yin, 1969; Farah et al., 1995), the composite-face effect (Young et al., 1987; Rossion and Boremanse, 2008; Mondloch and Maurer, 2008), the part-whole effect (Tanaka and Farah, 1993), the Thatcher illusion (Lewis and Johnston, 1997; Carbon et al., 2007), the N170 (Bentin et al., 1996; Carmel and Bentin, 2002), and activation in a brain region termed the fusiform face area (Kanwisher et al., 1997; Tong et al., 2000). These effects are attributed either partly or wholly, to the innate significance of faces (Slater and Kirby, 1998; Carmel and Bentin, 2002; Dekowska et al., 2008), to their inherent spatial properties (Diamond and Carey, 1986; McLaren, 1997), or to extensive experience with viewing faces in a particular orientation (Gauthier and Tarr, 1997; Gauthier et al., 1998; McLaren, 1997). To the extent that faces are treated preferentially *because* they are overlearned, the hallmarks of face-processing should be elicited after training for non-face objects. What is the role of learning in enabling objects to be treated preferentially by the brain, and how does it interact with the spatial properties of those objects?

4. Visual expertise, for example expertise in identifying breeds of dogs, is thought to engage a special type of processing: experts apprehend objects holistically, whereas novices analyze by parts (Tanaka and Farah, 1993; Bukach et al., 2006; Gauthier et al., 1998; Gauthier and Tarr, 2002; Busey and Vanderkolk, 2005). Expert object processing is the outcome of prolonged experience with - i.e., learning of - many examples of an object class. But it is also true that expertise can be limited within the object class to a subset of objects - to a particular breed of dogs, for example (Diamond and Carey, 1986), or to a particular race of faces (i.e., the other-race effect) (Valentine, 1991; Byatt and Rhodes, 2004). This implies that the generalization of expertise to novel conditions is constrained by the variation of examples used in training, and that the involvement of 'special' (holistic) processes might vary across exemplars. The study of learning of complex patterns can illuminate whether visual processing changes qualitatively after expertise, and how such a shift is brought about. It can also tell us how best to train people and robots to recognize, identify and classify images. There is practical value to this knowledge: accurate image classification is essential to many professional domains, such as medical diagnosis, surveillance and the fashion industry.

The approach taken in this dissertation is to address some of the above issues by

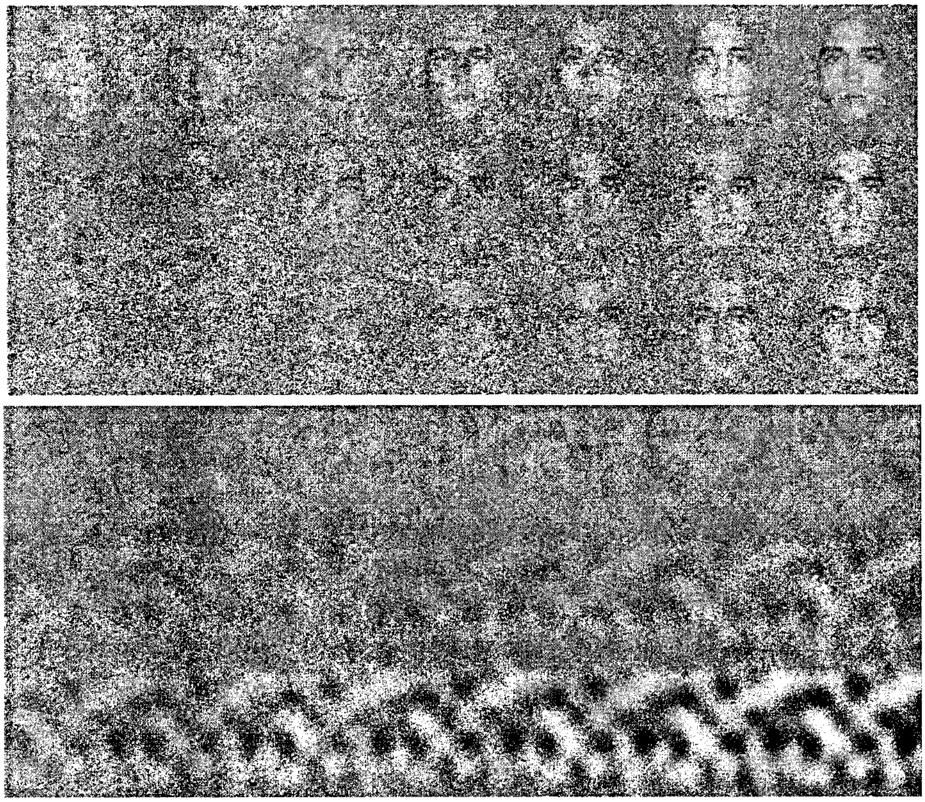


Figure 1.1: Examples of the stimulus conditions used in the experiments: a given face and texture, presented in each of three levels of external noise (increasing from top to bottom), at each of seven contrasts (increasing from left to right). The seven contrasts were unique at each noise level. Higher contrasts were used for the textures.

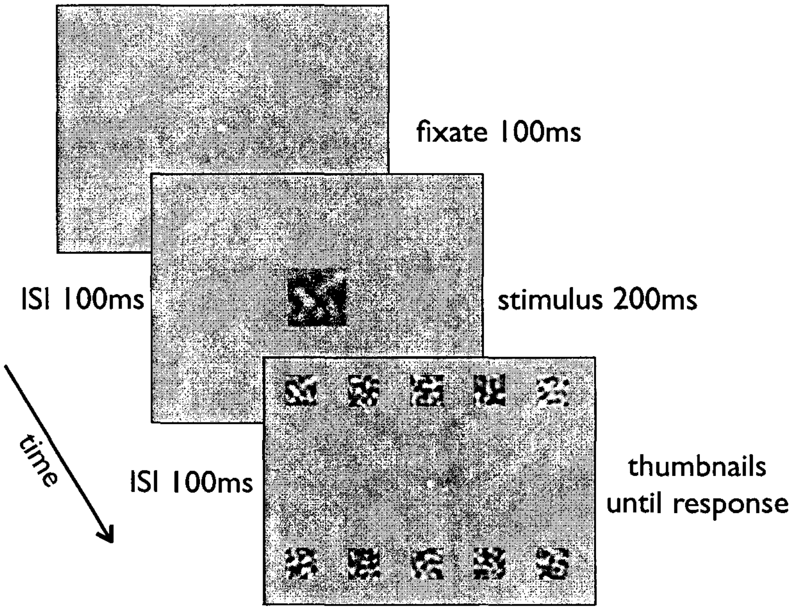


Figure 1.2: A schematic illustration of the trial sequence in the 10AFC identification task.

studying the learning of two kinds of complex patterns: faces, an object class with which we are already hugely familiar, and arbitrary texture patterns, with which we have no experience. Similar to other frequently encountered object classes, both faces and textures comprise multiple features. These stimuli afford a comparison of learning in already well-formed representations (faces), with representations that must be built from scratch (textures). Identification of both faces and textures does improve (Gold et al., 1999, 2004). When presented with a noisy face or texture, and asked to identify which of ten items it was, people reliably improve over time despite how good or bad they initially are at the task. The task involves extracting a signal from noise, and matching it to a noiseless exemplar, therefore the task challenges *perception*, and learning within this context is perceptual learning. In all the experiments reported, the signal (face or texture) is presented, in one of three levels of external noise (i.e., random perturbation of the luminance values of every stimulus pixel), at one of seven different contrast values. These 21 stimulus conditions are designed to probe the system across a range of difficulty levels (see Figure 1.1), and identification immediately post-presentation is designed to eliminate the effects of memory that might confound interpretation in a delayed recognition task (see Figure 1.2 for a schematic illustration of the task).

Chapter 2 (Inversion), identifies two ways in which learning of face- and texture identification is stimulus-specific, and relates this stimulus-specific learning to an index of face-specific processing known as the face-inversion effect. This chapter is in press in the journal Vision Research.

Chapter 3 (Contrast Reversal), looks at what in the image is learned, and identifies a third type of stimulus-specificity in learning of textures. This chapter is published in the Journal of Vision.

Chapter 4 (Sleep), examines the time-course of learning, and the contribution of sleep to the overall improvements on face identification. This chapter is published in Vision Research.

Chapter 5 (How much practice?), determines whether learning can be induced with small amounts of practice. This chapter is accepted pending revisions in Vision Research.

Chapter 6 (Retention), assesses whether the improvements on both tasks are intact a year later. This chapter is ready for submission.

Chapter 7 (Summary and the Future): as the title suggests.

Note that there is some redundancy in the introductions to each chapter, because these chapters were submitted for publication independently. In almost all chapters, the introduction will define perceptual learning and highlight some of the primary characteristics of learning. As a whole, these chapters cover the essential characteristics of learning of face- and texture identification. Learning of other complex visual tasks has been studied: visual search (Sireteanu and Rettenbach, 1995; Lobley and Walsh, 1998; Sigman and Gilbert, 2000), letter identification (Chung et al., 2006), object identification (Furmanski and Engel, 2000), and object categorization (Tanaka et al., 2005), but none have received the comprehensive treatment given to the tasks studied here.

References

Ball, K., Sekuler, R., 1982. A specific and enduring improvement in visual-motion discrimination. Science 218 (4573), 697–698.

Ball, K., Sekuler, R., 1987. Direction-specific improvement in motion discrimination. Vision Res 27 (6), 953–965.

Barlow, H., 1990. Conditions for versatile learning, Helmoltz's unconscious inference and the task of perception. Vision Research 30 (11), 1561–1571.

Bentin, S., Allison, T., Puce, A., Perez, E., McCarthy, G., 1996. Electrophysiological studies of face perception in humans. The Journal of Cognitive Neuroscience 8 (6), 551–565.

Bukach, C. M., Gauthier, I., Tarr, M. J., 2006. Beyond faces and modularity: the power of an expertise framework. Trends Cogn Sci 10 (4), 159–166.

Busey, T., Vanderkolk, J., 2005. Behavioural and electrophysiological evidence for configural processing in fingerprint experts. Vision Research 45, 431–448.

Byatt, G., Rhodes, G., 2004. Identification of own-race and other-race faces: implications for the representation of race in face space. Psychon Bull Rev 11 (4), 735–741.

Carbon, C., Gruter, T., Weber, J., Lueschow, A., 2007. Faces as objects of non-expertise: Processing of thatcherised faces in congenital prosopagnosia. Perception 36 (11), 1635–1645.

Carmel, D., Bentin, S., 2002. Domain specificity versus expertise: factors influencing distinct processing of faces. Cognition 83, 1–29.

Chung, S. T. L., Levi, D. M., Li, R. W., 2006. Learning to identify contrast-defined letters in peripheral vision. Vision Res 46 (6-7), 1038–1047.

Crist, R., Kapadia, M., Westheimer, G., Gilbert, C., 1997. Perceptual learning of spatial localization: Specificity for orientation, position and context. Journal of Neurophysiology 78 (6), 2889–2894.

Dekowska, M., Kuniecki, M., Jaskowski, P., 2008. Facing facts: Neuronal mechanisms of face perception. Acta Neurobiologiae Experimentalis 68 (2), 229–252.

Diamond, R., Carey, S., 1986. Why faces are and are not special: an effect of expertise. J Exp Psychol Gen 115 (2), 107–117.

Fahle, M., Edelman, S., Poggio, T., 1995. Fast perceptual learning in hyperacuity. Vision Res 35 (21), 3003–3013.

Farah, M. J., Tanaka, J. W., Drain, H. M., 1995. What causes the face-inversion effect? J Exp Psychol Human 21 (3), 628–634.

Fiorentini, A., Berardi, N., 1981. Learning in grating waveform discrimination: specificity for orientation and spatial frequency. Vision Research 21 (7), 1149–1158.

Furmanski, C. S., Engel, S. A., 2000. Perceptual learning in object recognition: object specificity and size invariance. Vision Res 40 (5), 473–484.

Gauthier, I., Tarr, M., 2002. Unraveling mechanisms for expert object recognition: Bridging brain activity and behaviour. Journal of Experimental Psychology: Human Perception and Performance 28 (2), 431–446.

Gauthier, I., Tarr, M. J., 1997. Becoming a "greeble" expert: exploring mechanisms for face recognition. Vision Res 37 (12), 1673–1682.

Gauthier, I., Williams, P., Tarr, M. J., Tanaka, J., 1998. Training 'greeble' experts: a framework for studying expert object recognition processes. Vision Res 38 (15-16), 2401–2428.

Gilbert, C., 1994. Early perceptual learning. Proceedings of the National Academy of Sciences of the Unites States of Americe 91, 1195–1197.

Gold, J., Bennett, P. J., Sekuler, A. B., 1999. Signal but not noise changes with perceptual learning. Nature 402 (6758), 176–178.

Gold, J. M., Sekuler, A. B., Bennett, P. J., 2004. Characterizing perceptual learning with external noise. Cognitive Sci 28, 167–207.

Kanwisher, N., McDermott, J., Chun, M., 1997. The fusiform face area: A module in human extrastriate cortex specialized for face perception. The Journal of Neuroscience 17 (11), 4302–4311.

Karni, A., Bertini, G., 1997. Learning perceptual skills: behavioral probes into adult cortical plasticity. Curr Opin Neurobiol 7 (4), 530–535.

Karni, A., Sagi, D., 1991. Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. Proc Natl Acad Sci U S A 88 (11), 4966–4970.

Karni, A., Sagi, D., 1993. The time course of learning a visual skill. Nature 365 (6443), 250–252.

Karni, A., Tanne, D., Rubenstein, B. S., Askenasy, J. J., Sagi, D., 1994. Dependence on rem sleep of overnight improvement of a perceptual skill. Science 265 (5172), 679–682.

Lewis, M., Johnston, R., 1997. The thatcher illusion as a test of configural disruption. Perception 26 (2), 225–227.

Lobley, K., Walsh, W., 1998. Perceptual learning in visual conjunction search. Perception 27 (10), 1245–1255.

McKee, S. P., Westheimer, G., 1978. Improvement in vernier acuity with practice. Perception and Psychophysics 24 (3), 258–262.

McLaren, I. P., 1997. Categorization and perceptual learning: an analogue of the face inversion effect. Q J Exp Psychol A 50 (2), 257–273.

Mondloch, C., Maurer, D., 2008. The effect of face orientation on holistic processing. Perception 37 (8), 1175–1186.

Rossion, B., Boremanse, A., 2008. Nonlinear relationship between holistic processing of indivdual faces and picture-plane rotation: Evidence from the face composite illusion. Journal of Vision 8 (4).

Schoups, A. A., Vogels, R., Orban, G. A., 1995. Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. J Physiol 483 (Pt 3), 797–810.

Sigman, M., Gilbert, C. D., 2000. Learning to find a shape. Nat Neurosci 3 (3), 264–269.

Sireteanu, R., Rettenbach, R., 1995. Perceptual learning in visual search: Fast, enduring but non-specific. Vision Research 35 (14), 2037–2043.

Slater, A., Kirby, R., 1998. Innate and learned perceptual abilities in the newborn infant. Experimental Brain Research 123 (1-2), 90–94.

Stickgold, R., Whidbee, D., Schirmer, B., Patel, V., Hobson, J. A., 2000. Visual discrimination task improvement: A multi-step process occurring during sleep. J Cogn Neurosci 12 (2), 246–254.

Tanaka, J., Farah, M., 1993. Parts and wholes in face recognition. Quarterly Journal of Experimental Psychology A- Human Experimental Psychology 46 (2), 225–245.

Tanaka, J. W., Curran, T., Sheinberg, D. L., 2005. The training and transfer of real-world perceptual expertise. Psychol Sci 16 (2), 145–151.

Tong, F., Nakayama, K., Moscovitch, M., Weinrib, O., Kanwisher, N., 2000. Response properties of the human fusiform area. Cognitive Neuropsychology 17 (1-3), 257–279.

Valentine, T., 1991. A unified account of the effects of distinctiveness, inversion, and race in face recognition. Q J Exp Psychol A 43 (2), 161–204.

Yin, R. K., 1969. Looking at upside-down faces. J Exp Psychol 81 (1), 141–145.

Young, A., Hellawell, D., Hay, D., 1987. Configurational information in face perception. Perception 16 (6), 747–759.

Chapter 2

Perceptual learning modifies inversion effects for faces and textures

2.1 Abstract

We trained subjects to identify either upright or inverted faces in a 10AFC task and measured performance subsequently in four conditions: same- and different-upright faces, and same- and different-inverted faces. Performance improved for both the upright-trained and the inverted-trained groups. The improvements were highly specific to the trained face exemplars, and largely specific to the trained face orientations. This pattern of results yielded an increase in the face-inversion effect after upright training, and a decrease in the inversion effect after inverted training, but only for the trained set of faces in both groups. A similar pattern of results was found for phase-scrambled faces in which the configural structure of faces had been removed: Although there was no baseline inversion effect for the scrambled stimuli, inversion effects emerged after training. We consider the implications of this pattern of learning for current views on the face-inversion effect, and face-encoding more generally.

Citation: Hussain, Z., Sekuler, A. B., & Bennett, P. J. Perceptual learning modifies inversion effects for faces and textures. *Vision Research*, In Press.

2.2 Introduction

Repeating a perceptual task improves the ability to detect, discriminate, and identify stimuli, a phenomenon known as perceptual learning (Ball and Sekuler, 1987; Fiorentini and Berardi, 1981; Rubin et al., 1997; Sigman and Gilbert, 2000; Yi et al., 2006). Often the benefits of perceptual learning are found only for the particular stimuli used during training. For example, practice improves sensitivity in a spatial frequency discrimination task, but the effects of practice are abolished by changing the target's spatial frequency by an octave, or its orientation by 90 deg (Fiorentini and Berardi, 1981). Similar specificity is found after training on visual tasks such as motion direction discrimination, contour perception, and figure-ground segmentation (Ball and Sekuler, 1987; Fiorentini and Berardi, 1981; Rubin et al., 1997; Sigman and Gilbert, 2000; Yi et al., 2006). The specificity of perceptual learning in these tasks has lead some researchers to suggest that the effects of learning alter the properties of low-level visual mechanisms (Karni and Bertini, 1997; Gilbert, 1994; Crist et al., 1997; Fahle, 2004).

Learning also occurs in more complex visual tasks. For example, accuracy in a face identification task improves significantly with practice (Dolan et al., 1997; Elliott et al., 1973; Gold et al., 1999b, 2004; Goldstein and Chance, 1985; McKone et al., 2007). Despite these demonstrations that laboratory-based practice improves performance, it generally is thought that a lifetime of perceiving faces has helped most human adults to become face identification experts. Indirect support for this view comes from the other-race effect, in which people are better at recognizing and identifying faces from their own racial group than faces from other groups (Valentine and Bruce, 1986; Tanaka et al., 2004; Byatt and Rhodes, 2004), and from many demonstrations that face identification is poorer for inverted faces than upright faces (Yin, 1969; Valentine, 1988). These effects can be interpreted as evidence for limited generalization of face expertise to unfamiliar exemplars and orientations, and they resemble, at least qualitatively, the stimulus-specific effects found in many studies of perceptual learning (Fiorentini and Berardi, 1981; Furmanski and Engel, 2000; Sigman and Gilbert, 2000). However, we know of no direct demonstration of exemplar- and orientation-specific effects of perceptual learning with faces. The current experiments examine whether such effects can be induced by practice.

A second goal of the current experiments is to investigate the stimulus conditions needed to produce orientation-specific effects of learning. McLaren (1997) theorized that perceptual learning produces orientation-specific effects only when the object class contains average, or prototypical, structure that resembles individual exemplars of that class (see also Diamond and Carey 1986). Frontal views of faces comprise an object class with such prototypical structure, because the average of a large set of faces resembles a typical face. Scrambling the phase spectra of a set of faces produces a set of textures that do not have prototypical structure, because the average of a large set of textures will be (approximately) a uniform field. Hence, McLaren's theory predicts that orientation-specific effects of learning should be obtained with faces but not textures. The current experiments tested this prediction.

This study consists of three experiments. The first examines whether learning of upright faces generalizes to a novel set of upright faces. The second experiment tests whether learning of upright or upside-down faces generalizes to faces that have been rotated by 180 deg. The third experiment compares the effects of learning obtained with faces to those obtained with textures. The results indicate that learning with faces is, in part, both exemplar- and orientation-specific, and that that the effects of learning are similar with faces and textures.

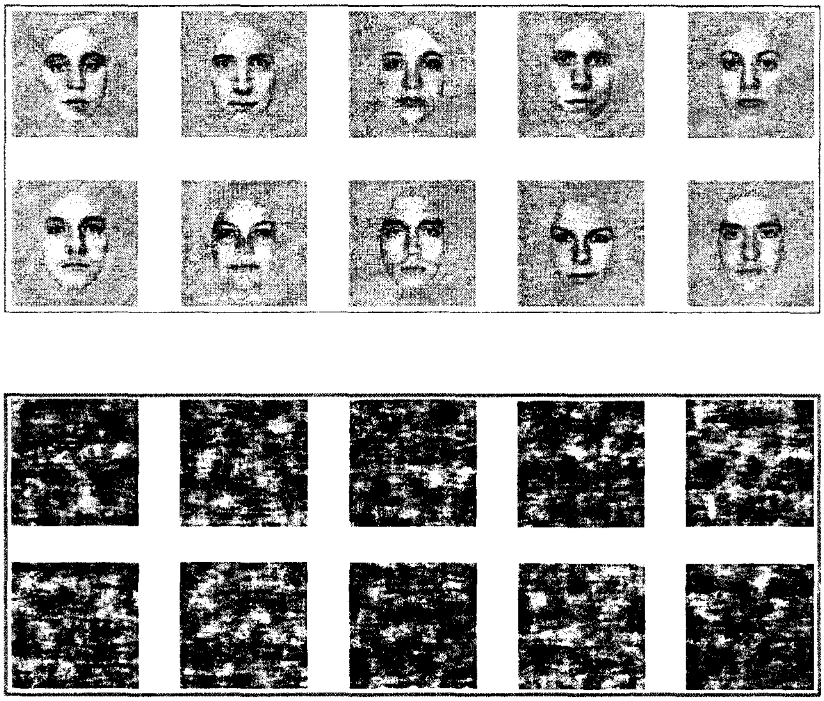


Figure 2.1: Examples of the face stimuli (Experiments 1 and 2), and the phase-scrambled stimuli (Experiment 3) used for the 10AFC identification task.

2.3 Methods

2.3.1 Subjects

Forty-seven subjects between the ages of 18-32 years (M = 19.9 years) took part either for remuneration ($10/hour) or for partial course-credit. All subjects had normal or corrected-to-normal visual acuity as measured by the Snellen acuity chart. Twenty-four subjects were in the same-face group, and 23 subjects were in the different-face group.

2.3.2 Apparatus and Stimuli

Stimuli were generated on a Power Mac G4 computer using Matlab (The Mathworks, version 5.2.1) and the Psychophysics and Video Toolboxes (Brainard, 1997; Pelli, 1997). They were displayed on a Sony Trinitron GDM-F520 monitor set to a resolution of 1024 x 768 pixels and a frame rate of 85 Hz (non-interlaced). Average luminance was 49 cd/m^2. The monitor calibration data were used to build a 1779-element lookup table (Tyler et al., 1992), and customized computer software constructed the stimuli on each trial by selecting the appropriate luminance values from the calibrated lookup table and storing them in the display's eight-bit lookup table.

The methods used to create the face stimuli have been described previously (Gold et al., 1999a). Twenty faces – 10 male and 10 female – were cropped to display only internal features within an oval subtending 190:140 pixels (subtending 3.6 x 2.6 deg at the viewing distance of 114 cm), and equated in terms of their amplitude spectra. Faces were presented in a square frame (256 x 256 pixels, or approximately 4.8 x 4.8 deg). The 20 faces were randomly divided into two sets (sets A and B), with the constraint that each set of ten faces comprised five male and five female faces (see Figure 2.1). During the experiment, stimulus contrast was varied across trials using the method of constant stimuli. Seven levels of contrast were spaced approximately equally on a logarithmic scale, and spanned a range that was sufficient to produce significant changes in performance in virtually all subjects (see Table 2.1). The images were shown in three levels of static two-dimensional Gaussian noise, created by sampling from distributions with contrast variances of .001, .01, and .1. Hence, there were a total of 21 stimulus conditions (seven contrast levels x three external noise levels) that allowed subjects to view each face at a variety of signal-to-noise ratios.

Table 2.1: RMS contrasts of the upright faces. Contrasts of the inverted faces and phase-scrambled stimuli used in Experiments 2 and 3 were 1.5 times greater than the values shown here.

Low Noise	Medium Noise	High Noise
.0049	.0154	.0487
.0059	.0186	.0226
.0071	.0226	.0715
.0087	.0273	.0866
.0105	.0332	.1049
.0127	.0401	.1269
.0154	.0487	.1539

2.3.3 Procedure

Subjects viewed the monitor binocularly from a distance of 114 cm. Viewing position was stabilized with a chin/forehead rest. The stimulus display was the only source of illumination in the testing room. A testing session began with a 60 s adaptation period during which time the subject viewed a uniform field set to the average luminance of the stimulus display. Following adaptation, each trial began with the presentation of a central fixation point for approximately 100 ms (black spot, 0.15 x 0.15 deg), followed by a randomly selected face presented for approximately 200 ms at the center of the screen. The stimulus conditions were intermixed, so the target face on each trial was selected from one of the 21 conditions (i.e., 7 contrasts x 3 noise levels). After the face disappeared, the entire set of 10 faces was presented as noiseless, high-contrast thumbnail images each subtending approximately 1.7 x 1.7 deg. Five thumbnails were presented on the top half of the screen, and five on the bottom half, and the location of specific face identities was constant across trials and across subjects. The subject's task was to decide which of the 10 faces had been presented during the trial, and to respond by clicking on the chosen face with the mouse. Auditory feedback was provided after each response (high- and low-pitched tones for correct and incorrect responses, respectively), and the next trial began one second after feedback.

All subjects participated in the experiment on two consecutive days. On Day 1, each subject performed the face identification task with one of the two sets of 10 faces. On Day 2, subjects in the same-face group performed the identification task with the same faces they saw on Day 1, but subjects in the different-face group performed the task with the set of 10 faces that they had not seen on Day 1. The order of sets was counterbalanced across subjects.

On both Days 1 and 2, subjects performed 40 trials per stimulus condition for a total of 840 trials (40 trials x 21 stimulus conditions), which were completed in approximately one hour. Each face was selected randomly (with replacement) on each trial, such that on average each face was shown approximately 84 times during the entire session.

2.4 Results

Among the statistical analyses described here and in subsequent Results sections are comparisons of i) average performance on Days 1 and 2; ii) performance in different bins of trials on Day 1; and iii) performance in different bins of trials on Day 2. Many of these comparisons are mutually orthogonal, and therefore provide independent estimates of the effects of practice on performance. Furthermore, the effect of learning measured across days depends in part on the definition of baseline performance. Instead of selecting one baseline arbitrarily, we have used different definitions of baseline performance to reveal different aspects of learning. In particular, it will be shown that using different baselines provides different estimates of the amount and generalization of learning.

For the purpose of the analyses, the 840 trials on each day were divided into eight blocks of 105 sequential trials (trial bins 1-8). For each bin, the proportion of correct responses was calculated after collapsing across all levels of stimulus contrasts and noise. Proportion correct at each bin within the session on Day 1 and Day 2 is plotted for both groups in Figure 2.2. On Day 1, performance increased across bins but was similar in the both groups: a 2 (Group) x 8 (Bin) ANOVA found a significant effect of Bin ($F(7,315) = 62.2$, $p < .0001$), but the main effect of Group ($F(1,45) = 0.38$, $p = .54$) and the Bin x Group interaction ($F(7,315) = .89$, $p = .51$) were not significant. The lack of an interaction suggests that accuracy improved at similar rates in both groups. On Day 2, there also was a main effect of Bin ($F(7,315) = 24.89$, $p < .0001$), which indicates that performance generally improved during the session. However, unlike what was found on Day 1, there was a significant effect of Group ($F(1,45) = 11.01$, $p = .002$), indicating that response accuracy was lower in the different-face group, and a significant Bin x Group interaction ($F(7,315) = 6.56$, $p < .0001$), indicating that the increase in accuracy during Day 2 was greater in the different-face group than in the same-face group.

Overall proportion correct was calculated for each group by collapsing responses across all levels of stimulus contrast and external noise. In the same-face group, overall proportion correct was 16% higher on Day 2 than Day 1, a difference that was statistically significant ($t(23) = 11.09$, $p < .0001$, one-tailed). In the different-face group, proportion correct also was 4% higher on Day 2, a difference that was significantly greater than zero ($t(22) = 1.92, p = .034$, one-tailed) but significantly less than the improvement in accuracy attained by the same-face group ($t(45) = 5.34, p < .0001$, one-tailed). A slightly different perspective on the between-day effect is gained by comparing performance in Bins 8 and 9. In the same-face group, accuracy in Bin 9 was 5% higher than accuracy in Bin 8. In the different-face group, on the other hand, accuracy was 14% lower in Bin 9 than in Bin 8. This difference between groups was confirmed by a t-test on the difference scores between Bins 8 and 9 ($t(45) = 5.420, p < .0001$). These analyses suggest that the performance of both groups improved across days, but that the improvement was significantly greater in the same-face group.

We calculated the difference in accuracy between Bin 9 and Bin 1 for both groups. Performance in these bins represents initial performance on Days 1 and 2. The same group was 29% more accurate in Bin 9 than in Bin 1 ($t(23) = 13.83, p < .0001$), and the different group was 7% more accurate in Bin 9 than in Bin 1 ($t(22) = 13.83, p < .0001$). The group difference in the difference scores between Bins 9 and 1 was also significant ($t(44) = 6.71, p < .0001$). The 7% improvement in Bin 9 relative to Bin 1 shown by the different face group can be thought of as the component of learning that is not stimulus specific. We will refer to this effect as the task-general component of learning

2.5 Discussion

The group differences in performance on Day 2 make clear that the same-face group was at an advantage compared to the different-face group. Improvement across sessions was greater when the same stimuli were viewed on both days, and the different-face group, despite greater amounts of learning on Day 2 did not, on average, achieve the level of performance of the same-face group. Thus, the major part of what was learned did not transfer across stimulus sets, but a small proportion (7%) did. This transfer was indicated by the increased accuracy for the different face group in Bin 9 relative to Bin 1, which we consider the task-general component of learning.

Although previous work has shown that face identification can be improved by practice (Dolan et al., 1997; Elliott et al., 1973; Gold et al., 1999b; Goldstein and Chance, 1985; McKone et al., 2007), exemplar-specific improvements in face perception have not

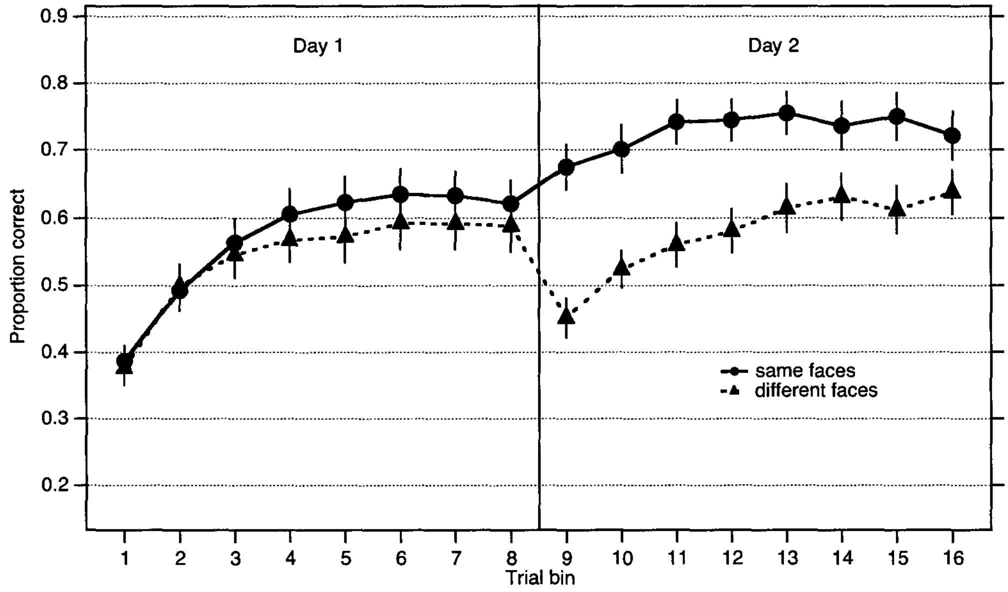


Figure 2.2: The results of Experiment 1. Proportion correct of responses, collapsed across levels of stimulus contrast and noise, calculated for successive bins of 105 trials on Days 1 and 2. Solid traces represent performance of the group ($N = 24$) that did the task with the same faces on both days. Dashed traces represent performance of the group ($N = 23$) that did the task with different faces on Day 1 and Day 2.

been reported previously. The perceptual learning literature is replete with examples of stimulus-specificity for tasks involving discriminations along a single stimulus attribute, such as spatial frequency or motion direction (Ball and Sekuler, 1987; Fiorentini and Berardi, 1981). In these cases, learning effects are minimized or abolished after changes to, for example, the trained motion direction or the trained stimulus orientation, suggesting plasticity at early visual sites that code basic stimulus properties (Karni and Bertini, 1997; for a different interpretation, see Mollon and Danilova, 1996). For complex perceptual tasks, learning is thought to modify sites where stimulus representations are transformation-invariant (Karni and Bertini, 1997), consistent with the study finding generalization of learning across face viewpoints (Moses et al., 1996). Nevertheless, learning in some complex visual tasks does exhibit stimulus-specificity: For example, there is no transfer of learning from contrast- to luminance-defined letters during letter identification, across shapes in a figure-ground segmentation task, or across object identity in an object identification task (Chung et al., 2006; Furmanski and Engel, 2000; Sigman and Gilbert, 2000). And recent work has found evidence for face-specific adaptation. For example, face adaptation aftereffects only occur along the geometric trajectory of the exposed face identity (Leopold et al., 2001), and these distortions do not transfer across viewpoint (Anderson and Wilson, 2005; Jeffery et al., 2006). In our experiment, we found relatively little transfer of learning to novel face exemplars. Thus, the present results support stimulus-specificity as a characteristic of learning even for complex tasks, such as face-identification, with which we have expertise.

2.6 Experiment 2

The goal of Experiment 2 was to determine if the learning shown in Experiment 1 was orientation-specific. Specifically, we asked whether learning transfers from faces trained in the upright orientation to the same faces viewed in the inverted orientation, and vice versa. Earlier work showing some transfer of learning across face viewpoints (Moses et al., 1996), indicates some potential for learning to generalize beyond the particular images viewed during training. And, a recent study using a set of houses with similar spatial configurations, showed that learning partially generalized to untrained orientations, suggesting that with such stimuli some proportion of learning is orientation invariant (Husk et al., 2007). On the other hand, if learning of faces is like most other low-level tasks, then the effects of practice should be specific to the trained orientations. Additionally, we consider the consequences of specificity of learning for the size of the face-inversion effect.

2.7 Methods

2.7.1 Subjects

Thirty-two subjects (18 female) between the ages of 18-32 years ($M = 20.5$ years) participated either for partial course credit or remuneration ($10/hour). All had normal or corrected-to-normal visual acuity as measured by the Snellen acuity chart, and none had participated in Experiment 1. Sixteen subjects were assigned to the upright-training group, and 16 to the inverted-training group.

2.7.2 Apparatus and Stimuli

The apparatus and two sets of faces were the same as in Experiment 1. The inverted faces were displayed at higher contrasts than the upright faces (see Table 1) because pilot work showed that performance on inverted-face identification with the contrasts used for the upright faces was below chance (consistent with the face-inversion effect: Farah et al., 1995; Friere et al., 2000; Tanaka et al., 2004; Valentine and Bruce, 1986; Yin, 1969; Sekuler et al., 2004; Gaspar et al., 2008a). The contrast variances used for the upright faces at each noise level are shown in Table 1.Contrast variances of the inverted faces were 1.5 times those of the upright faces.

2.7.3 Procedure

All subjects participated in the experiment on two consecutive days. On Day 1, subjects were randomly assigned to train on the face identification task with one set of faces either in the upright or inverted orientation. The task protocol was the same as described for Experiment 1. On Day 1, subjects performed 40 trials per stimulus condition for a total of 840 trials at the training orientation. On Day 2, all subjects performed the identification task with both sets of faces in both orientations. Thus, all subjects performed the task in four experimental conditions on Day 2, and only one of those conditions was identical to training Day 1. The four experimental conditions (upright/inverted x same/different) were blocked on Day 2, with the order of blocks counterbalanced across subjects. For each experimental condition, subjects performed 10 trials per stimulus condition for a total of 210 trials per block. Thus, all subjects performed a total of 840 trials across the entire session (210 trials/block x 4 blocks).

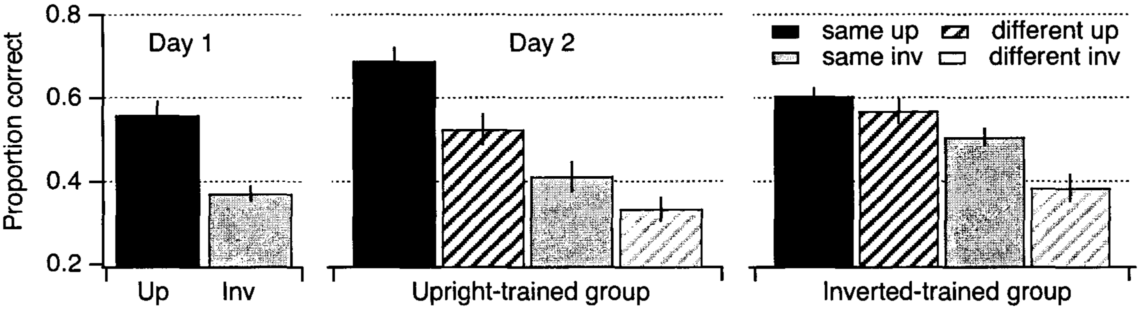


Figure 2.3: Identification performance (overall proportion correct) on Day 1 and Day 2 for faces measured in Experiment 2. Day 1: upright-trained (dark) versus inverted-trained (light) group. Day 2: upright versus inverted faces (dark versus light bars); same versus different faces (solid versus striped bars).

2.8 Results

2.8.0.1 Overall accuracy: Upright-trained faces

We first consider the effects of learning obtained with the group of subjects who were trained with upright faces on Day 1 (i.e., the upright-trained group; dark bar in the leftmost panel, and all the bars in the middle panel, Figure 2.3). We defined baseline performance as response accuracy measured on Day 1. Proportion correct in the same-upright condition on Day 2 was 13% greater than baseline performance, indicating that performance improved across days ($t(15) = 8.67$, $p < .0001$). Performance with different-upright faces on Day 2 did not differ from baseline performance ($t(15) = 1.39$, $p = .18$), and was 16% lower than accuracy obtained with same-upright faces on Day 2 ($t(15) = 8.62$, $p < .0001$). These results show that the effects of practice with upright faces on Day 1 did not generalize to novel upright faces on Day 2. Response accuracy measured on Day 2 with the same faces viewed upside-down (i.e., the same-inverted condition) and with novel inverted faces (i.e., the different-inverted condition), did not differ significantly from performance measured with inverted faces in a different group of subjects on Day 1 (same-inverted condition: $t(30) = .95$, $p = .35$; different-inverted condition: $t(30) = 1.05$, $p = .30$), which indicates that learning was orientation-specific. However, response accuracy on Day 2 was significantly greater in the same-inverted condition than in the different-inverted condition (8% advantage, $t(15) = 3.54$, $p = .003$), which suggests that within the familiar set of faces, some benefits of practice with upright faces on Day 1 did generalize to the opposite orientation on Day 2.

2.8.0.2 Inverted-trained faces

Next, we consider the effects of learning obtained with the group of subjects who were trained with inverted faces on Day 1 (i.e., the inverted-trained group; light bar in the leftmost panel, and all bars in the rightmost panel, Figure 2.3). As before, baseline performance was defined as response accuracy measured on Day 1. Proportion correct in the same-inverted condition on Day 2 was 13% better than baseline performance ($t(15) = 9.39$, $p < .0001$), indicating that performance improved across days. Performance with different-inverted faces on Day 2 did not differ from baseline performance ($t(15) = .481$, $p = .64$), and was 12% lower than accuracy with same-inverted faces on Day 2 ($t(15) = 4.12$, $p < .001$). These results show that, as was found with upright faces, the effect of practice with inverted faces did not generalize to novel inverted faces. Response accuracy measured on Day 2 with the same faces viewed in a new, upright orientation (i.e., the same-upright condition) and with novel, upright faces (i.e., the different-upright condition) did not differ significantly from baseline performance measured with upright faces in a different group on Day 1 (same-upright: $t(30) = 1.14, p = .26$; different-upright: $t(30) = .19, p = .85$), which shows that learning of inverted faces was orientation-specific, as was the case with upright faces. The difference between response accuracy in the same-upright and different-upright conditions on Day 2 was numerically less than that found in the complementary conditions in the upright-trained group, only approached significance (4% difference, $t(15) = 1.70, p = .10$).

We tested whether the amount of transfer across orientations was greater from the upright to inverted orientation than vice versa by comparing the difference between the same- and different-faces viewed in the untrained orientation in the two groups of subjects: the 8% advantage reported above for the upright-trained group versus the 4% advantage reported for the inverted-trained group. This comparison was not significant ($t(30) = 1.33$, $p = .19$), which suggests that there was no obvious advantage in transferring learning from upright to inverted stimuli, or vice versa.

2.8.1 Time-course of learning

We calculated proportion correct in eight consecutive bins of 105 trials on Day 1. On Day 2, each condition was separated into two bins – A and B – of 105 trials. Figure 2.4 shows that proportion correct improved on Day 1 for both tasks. From the first bin to the final bin on Day 1, proportion correct for upright face identification increased by 22% ($t(15) = 5.92$, $p < .0001$ one-tailed), and for inverted face identification it increased by

20% ($t(15) = 5.66$, $p < .0001$ one-tailed). Clearly, there was substantial within-session learning on Day 1 for both groups. For the following analyses, performance in Bin 1 on Day 1, which represents completely naive performance, is treated as baseline. As the analyses show, using this measure as baseline reveals transfer of learning on Day 2 that was not evident in the average response accuracy reported in the previous sections.

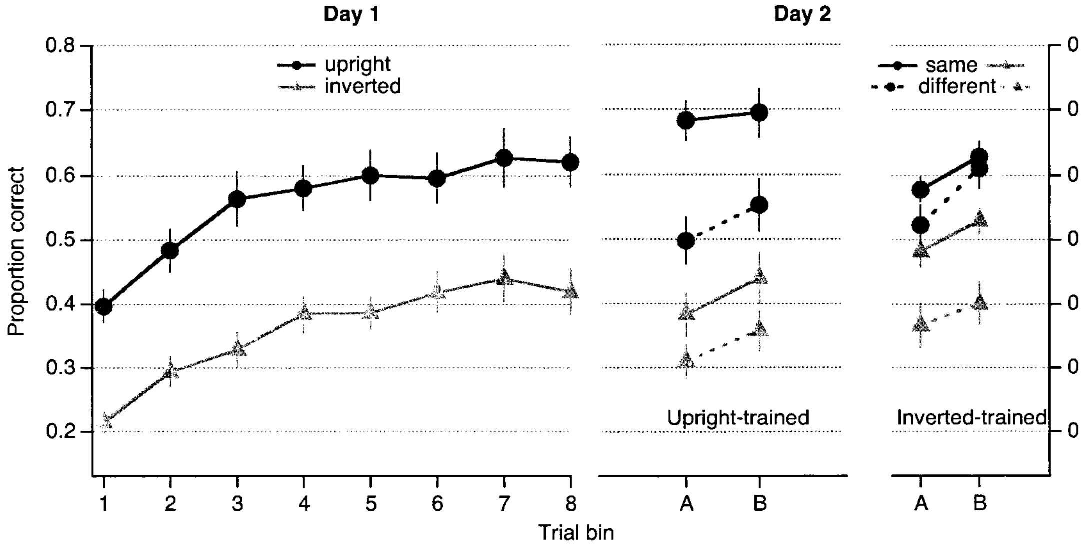


Figure 2.4: Time-course of learning for faces in Experiment 2. Left panel shows performance on Day 1. Right panels show performance in all four conditions on Day 2 separately for each group. Dark vs. Light: Upright vs. inverted stimuli; Solid vs. Dashed: Same vs. Different stimuli as those shown on Day 1. Each bin comprised 105 trials.

The upright-trained group's initial performance on Day 2 (Bin A) in the same-upright and different-upright conditions was, respectively, 29% better than baseline ($t(15) = 10.45$, $p < .00001$), and 10% better than baseline ($t(15) = 3.57$, $p = .0014$). Therefore, relative to completely naive performance, there was some transfer of learning to novel exemplars. Initial performance of the upright-trained group with inverted faces on Day 2 indicates that there was also some transfer across orientation: Proportion correct for same-inverted faces measured in Bin A was 16% higher than the baseline measured in Bin 1 from the inverted-trained group ($t(30) = 4.12$, $p = .0004$), and proportion correct for different inverted faces in Bin A was 9% better than baseline ($t(30) = 2.79$, $p = .01$). These comparisons indicate that the effects of training with upright faces partially transferred to inverted faces. Additionally, performance with the same-inverted faces in Bin A was 7% better than with different-inverted faces ($t(15) = 3.26$, $p = .005$), which reflects the effect of familiarity at the untrained orientation, over and above the task-general advantage that we describe in Experiment 1.

The upright-trained group's performance in Bin A was compared to its performance in Bin 8 the previous day (i.e., performance at the end of Day 1). Accuracy in the same-upright and different-upright conditions in Bin A were respectively 6% better than Bin 8 ($t(15) = 2.96, p = .009$) and 12% *worse* than Bin 8 ($t(15) = 3.76, p = .001$). Therefore, relative to accuracy at the end of the session on Day 1, there was a drop in performance with novel exemplars, whereas performance with the same exemplars improved. Proportion correct in Bin 8 did not differ from performance in Bin A in the same-inverted condition ($t(28.751) = .96, p = 0.34$), but was 12% higher than performance in the different-inverted condition ($t(29.625) = 2.75, p = .009$). These results are consistent with those described in the previous paragraph, and indicate that the effects of training with upright faces partially transferred to the same inverted faces but not to different inverted faces (i.e., over and above the task-general transfer indicated by comparison to bin 1, preceding paragraph).

The rightmost panel of Figure 2.4 shows performance of the inverted-trained group. For this group, initial performance on Day 2 in the trained condition (same-inverted, Bin A) was 26% higher than the Bin 1 baseline ($t(15) = 9.166, p < .0001$), which is equivalent to the amount learned in the trained condition by the upright-trained group. Initial performance with different-inverted faces on Day 2 was 15% better than baseline $t(15) = 4.67$, $p = .0003$), indicating that some learning transferred to novel stimuli in the same orientation, as was the case with the upright-trained group. There was also some transfer across orientation: performance in Bin A with same upright faces on Day 2 was 18% better than the Bin 1 baseline measured from the upright-trained group ($t(30) = 5.4$, $p < .0001$), and performance with different upright faces was about 13% better than the Bin 1 baseline ($t(30) = 3.05$, $p = .004$), indicating that, as was the case with the upright-trained group, learning transferred to the opposite orientation. Also, performance with same-upright faces in Bin A was 5% higher than performance with different-upright faces, a difference which approached significance ($t(15) = 1.97$, $p = .068$), and which reflects the effect of familiarity at the untrained orientation, as was found with the upright-trained group.

For the inverted-trained group, proportion correct in the same-inverted and different-inverted conditions in Bin A were respectively, 5% better than in Bin 8 ($t(15) = 2.322, p = .03$), and no different than in Bin 8 ($t(15) = 1.5, p = .14$). Therefore, performance in conditions using inverted faces improved relative to Bin 8 only when the faces were the same ones seen on Day 1. Proportion correct in the same-upright condition in Bin A was not significantly different than performance of the upright-trained group in Bin 8 ($t(22.34) = .96, p = .34$), and proportion correct in the different-upright condition was 10% worse than performance of the upright-trained group in Bin 8, a difference than approached significance ($t(28.59) = 1.9, p = .06$). These two comparisons suggest that practice with inverted faces on Day 1 generalized to the same faces presented upright, and are consistent with the effects of familiarity at the untrained orientation reported in the preceding paragraph.

Overall, transfer of learning across orientations was greater in the same-face condition than in the different-face condition. Additionally, we found that comparisons of average accuracy (Figure 2.3) were less sensitive tests of generalization of learning across exemplars and across orientations. Generalization becomes apparent when Day 2 performance is compared to performance during the first 105 trials on Day 1 (Figure 2.4). The transfer effects found with novel faces at the trained orientation are similar to the small amount of transfer found across stimulus sets in Experiment 1 in the comparison between initial trial bins on each day, which we attributed to task-general learning. The advantage of trained over novel faces at the untrained orientations reflects additional transfer beyond task-general transfer, i.e., it reflects transfer of learning across orientation.

2.8.1.1 Inversion effect

The inversion effect on Day 1, or baseline inversion effect, was defined as the between-group difference in response accuracy. Performance with upright faces was 19% better than performance with inverted faces on Day 1 ($t(30) = 4.76$, $p < .0001$, one-tailed). This difference in proportion correct occurred even though stimulus contrast was higher for inverted faces. Therefore, this inversion effect, although quite large, actually *underestimates* the inversion effect that would have been obtained had the stimulus contrasts been identical in both orientations.

On Day 2, the inversion effect was calculated within each group, separately for the trained and novel sets (i.e., the difference in proportion correct between the upright and inverted conditions, for the trained and novel sets). The inversion effects before and after training, for both training groups, are shown in Figure 2.5. Separate t tests were used to compare each inversion effect measured on Day 2 to the baseline inversion effect. For the upright-trained group, the size of the inversion effect in the same-face condition increased

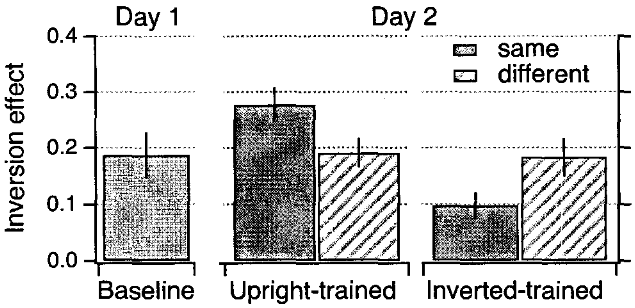


Figure 2.5: Face-inversion effect (upright minus inverted) measured in Experiment 2. The Baseline inversion effect measured on Day 1 was calculated by comparing performance of different groups of subjects; the error bar represents the standard error of the difference between means. On Day 2 the inversion effect was calculated by comparing performance within each subject in the upright-trained and inverted-trained groups. Error bars on Day 2 represent ±1 standard error of the mean.

significantly by 8% relative to baseline due to improved performance with the same upright faces ($t(15) = 2.79$, $p = .01$). However, the inversion effect in the different-face condition on Day 2 did not differ from baseline ($t(15) = .02, p = .97$). For the inverted-trained group, the inversion effect in the same-face condition decreased significantly by 10% relative to baseline due to improved performance with the trained upside-down faces $t(15) = 3.93$, $p = .001$). As with the upright-trained group, the inversion effect measured in the different-face condition did not differ from baseline ($t(15) = .19$, $p = .84$). Hence, practicing with upright faces on Day 1 produced a greater inversion effect on Day 2, but practicing with inverted faces on Day 1 produced a *smaller* inversion effect on Day 2. The magnitude of the learning effect was approximately equal (though of opposite sign) in the two groups, and was restricted to the trained faces.

2.9 Discussion

Experiment 2 showed that, in addition to being exemplar-specific, perceptual learning of faces is largely, though not entirely, orientation-specific, regardless of whether faces were trained in the upright or upside-down orientation. These instances of specificity are similar to those found with low-level tasks, suggesting that specificity is a general principle of learning across a range of tasks. The results also show clearly that the face-inversion effect can be modified selectively with experience. In the current experiment, practice improved the identification of inverted faces by the same amount as identification of upright faces. Although prior research has shown that practice improves the identification of upright faces (Dolan et al., 1997; Elliott et al., 1973; Gold et al., 1999b; Goldstein and Chance, 1985; McKone et al., 2007) and inverted faces (Robbins and McKone, 2003), this is the first demonstration that the effect of familiarity on the inversion effect is confined to the exposed set of faces. The pattern of results we find is at odds with an early report of no effect of familiarity on the inversion-effect (Scapinello and Yarmey, 1970), which could be accounted for by methodological differences. Scapinello and Yarmey (1970) measured errors made in an old-new recognition paradigm, while manipulating the number of item exposures; additionally, the inverted stimuli were not exposed during the study phase. In the current study, it is noteworthy that although the inversion effect decreased for the inverted-trained group, performance with upright faces on average was unaffected for this group. Thus, the size of the inversion effect can be altered by changes in the way inverted faces alone are encoded and/or represented.

2.10 Experiment 3

Experiment 2 found evidence for partial transfer of learning across orientation: on Day 2, performance with familiar stimuli shown in the opposite orientation was better than performance with novel, inverted stimuli. Additionally, when completely naive performance (Bin 1) was treated as baseline, both groups showed generalization of learning to novel exemplars and to the opposite orientation.

All faces share the same first-order structure (eyes-over-nose-over-mouth), which confers on them a clearly distinguishable canonical orientation. This type of structure has also been described as prototypical or average, wherein the pixel-wise average of a set of face exemplars could itself be considered a member of that set (McLaren, 1997). It is possible that transfer across orientation was facilitated by the presence of such structure in the stimuli, which is consistent with the transfer of learning across orientation found with houses but not band-limited textures by Husk et al. (2007). The houses used by Husk et al (2007), shared the same first-order structure, however the band-limited textures did not. Additionally, it has been proposed that inversion effects are contingent on prototypical structure within a stimuli class (McLaren, 1997). McLaren (1997) showed that inversion effects arise after familiarization with checkerboards that contain average structure, whereas checkerboards without average structure do not yield inversion effects. In Experiment 3 we examined inversion effects and transfer of learning for patterns in which the structure present in faces is diminished by virtue of phase-scrambling. Phase-scrambling removes the first-order structure present in faces, however it preserves the spatial frequency content of faces (Figure 1). The design used was identical to Experiment 2.

2.11 Methods

2.11.1 Subjects

Forty-eight subjects (12 males; 36 females) between the ages of 18-32 years ($M = 20.5$ years) participated in the experiment either for partial course credit or remuneration ($10/hour). None of the subjects had participated in the previous experiments. All had normal or corrected-to-normal Snellen acuity. Twenty-four subjects were assigned to the upright-training group, and 24 to the inverted-training group.

2.11.2 Apparatus and Stimuli

The apparatus was the same as the one used in Experiments 1 and 2. Two sets of phase-scrambled images were created from the sets of faces used in Experiments 1 and 2 by combining the average Fourier amplitude spectrum of the faces with 20 phase spectra derived from different samples of white, Gaussian noise. These stimuli have no canonical orientation, therefore upright and inverted were arbitrary labels for the presented orientations. Stimulus contrasts were the same as those used for the inverted faces in Experiment 2.

2.11.3 Procedure

The procedure was identical to that in Experiment 2 except that subjects were told that the task was pattern identification rather than face identification.

2.12 Results

2.12.1 Overall accuracy

Proportion correct for this experiment, calculated exactly as was done for Experiment 2, is shown in Figure 2.6. Performance for upright and inverted stimuli was the same on Day 1 ($t(46) = 0.6, p = .55$), indicating that there was no baseline inversion effect consistent with the fact that these stimuli did not have a canonical orientation.

We first consider the effects of practice obtained with subjects who were trained with so-called upright textures on Day 1 (i.e., the upright-trained group; Figure 2.6 middle panel). Proportion correct for same-upright textures on Day 2 was approximately 10% higher than baseline (i.e., performance on Day 1), indicating that there was significant learning in this condition ($t(23) = 6.7$, $p < .0001$). Performance in the different-upright condition on Day 2, however, was no different than baseline ($t(23) = .63$, $p = .54$), and was 11% lower than accuracy in the same-upright condition on Day 2 ($t(23) = 3.79$, $p = .0005$). These results indicate that the effects of practice did not generalize to novel textures. Accuracy in the same-inverted condition and the different-inverted condition on Day 2 did not differ from baseline accuracy measured in the inverted-trained group on Day 1 (same-inverted: $t(46) = .59$, $p = .56$; different-inverted: $t(46) = .68$, $p = .50$), which indicates that the effects of learning did not generalize to the opposite orientation. Additionally, there was no difference between accuracy for same-inverted and different-inverted textures ($t(23) = 0$, $p = 1$), which indicates that unlike what was found with faces, there was no transfer of learning across orientation within the familiar stimulus set.

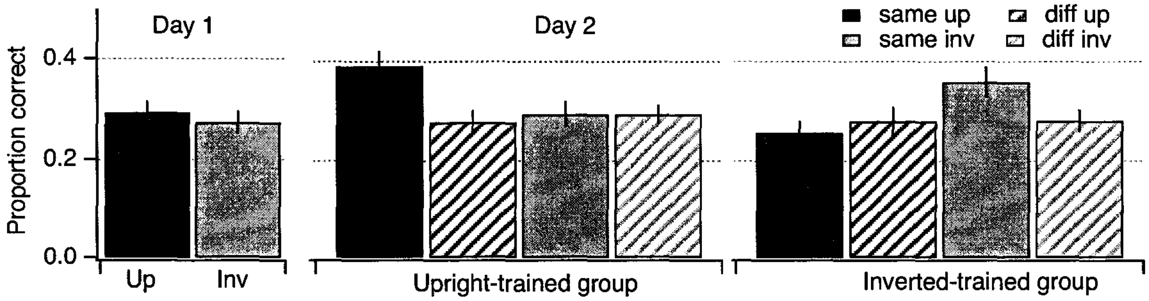


Figure 2.6: Identification performance (overall proportion correct) on Day 1 and Day 2 for phase-scrambled faces measured in Experiment 3. Note that the textures did not have a canonical orientation, so classification as upright and inverted was arbitrary. Day 1: upright-trained (dark) versus inverted-trained (light) group. Day 2: upright versus inverted faces (dark versus light bars); same versus different faces (solid versus striped bars).

Next, we consider the results obtained from subjects who were trained with so-called inverted textures on Day 1 (i.e., the inverted-trained group; Figure 2.6, right-most panel). Given that there was no effect of orientation on Day 1, it was not surprising to find that the pattern of results obtained from the inverted-trained group was the same as that obtained from the upright-trained group. Performance in the same-inverted condition on Day 2 was 8% higher than baseline ($t(23) = 5.71$, $p < .0001$), indicating that learning occurred. However, proportion correct in the different-inverted condition on Day 2 did not differ from baseline ($t(23) = 0.49$, $p = .63$), and was 8% lower than in the same-inverted condition on Day 2 ($t(23) = 4.89$, $p < .0001$). Both of these comparisons demonstrate that the effects of practice did not generalize to new textures. As with the upright-trained textures, there was no generalization of learning to the opposite orientation. Accuracy in the same-upright condition, and the different-upright condition on Day 2 did not differ from baseline accuracy measured with the upright-trained group on Day 1 (same-upright: $t(46) = 1.1, p = .28$; different-upright: $t(46) = .35, p = .73$). Also, there was no difference in accuracy for the same-upright and different-upright textures for this group on Day 2 ($t(23) = 1.03, p = .32$), indicating no transfer of learning across orientation within the familiar set of textures.

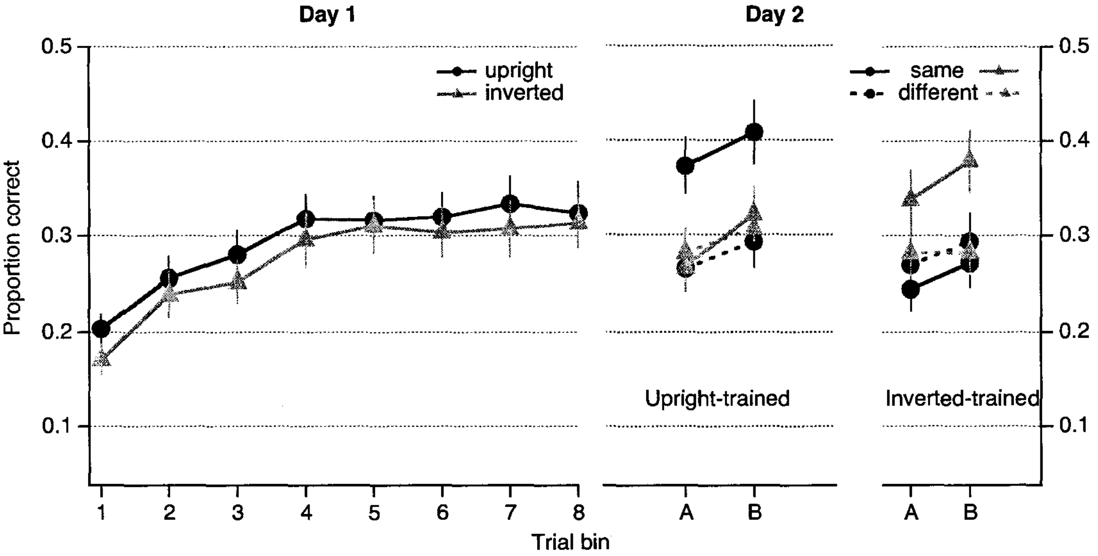


Figure 2.7: Time-course of learning for phase-scrambled faces in Experiment 3. Left panel shows performance on Day 1. Right panels show performance in all four conditions on Day 2 separately for each group. Dark vs. Light: Upright vs. inverted stimuli; Solid vs. Dashed: Same vs. Different stimuli as those shown on Day 1.

2.12.2 Time-course of learning

Figure 2.7 shows within-session performance on Days 1 and 2 for both groups, calculated as in Experiment 2. The traces on Day 1 confirm the absence of any orientation bias in the stimuli. In both groups, response accuracy increased significantly from the start to the end of the Day 1 session: by 12% in the upright-trained group ($t(23) = 3.89$,

$p < .0001$), and 14% in the inverted-trained group ($t(23) = 7.85$, $p < .0001$). The middle and right-hand panels of Figure 2.7 ~~shows~~ performance on Day 2, separately for each group, with each of the four conditions represented as in Figure 2.4. As was the case in Experiment 2, the time-course of learning reveals some generalization that could not be discerned in the average data. For the analyses below, performance in Bin 1 is considered baseline, as was done for time-course analyses in Experiment 2.

In the upright-trained group, initial performance on Day 2 (Bin A) in the same-upright condition was 17% better than baseline ($t(23) = 6.25, p < .0001$). Performance in Bin A did not differ across conditions that used novel stimuli (i.e., different-upright, same-inverted, different-inverted; $F(2, 46) = .336$, $p = .72$). Therefore, the average performance in these three conditions in Bin A was compared to baseline performance to assess whether there was any generalization of learning. A t test indicated that average performance in the untrained conditions was 7% better than the baseline measured from this group ($t(23) = 3.27$, $p = .004$), indicating that some of the benefits of learning transferred to the untrained exemplars and orientations. However, performance in Bin A in the same-upright condition was 10% better than performance in Bin A of the untrained conditions combined ($t(23) = 4.67$, $p < .0001$), so although some task-general advantage transferred, a large proportion of what was learned was exemplar-specific.

Proportion correct in Bin A in the same-upright condition was 5% higher than it was in Bin 8 the previous day (t(23) $= 2.98, p = .006$), whereas proportion correct in the different-upright condition in Bin A was 5% lower than in Bin 8, a difference that approached significance ($t(23) = 1.69, p = .10$). These comparisons show that there was a setback in learning relative to performance in Bin 8, when novel exemplars were used. Since performance in Bin A did not differ in the different-upright, same-inverted and different-inverted conditions, it can be inferred that learning also did not transfer to the opposite orientation. Exactly the same pattern of results was found with the inverted trained group.

The right-most panel of Figure 2.7 shows the performance of the inverted-trained group, whose pattern of performance was identical to the upright-trained group. In the inverted-trained group, performance in the same-inverted condition in Bin A was 17% better than baseline ($t(23) = 6.881, p < .0001$). In Bin A, there was no difference in performance among conditions that used novel stimuli ($F(2, 46) = 1.86$, $p = .17$), so the average of those three conditions was compared to baseline to test for generalization of learning. The average was 9% greater than baseline, a difference that was statistically significant ($t(23) = 5.32$, $p < .0001$), which shows that there was some generalization of learning relative to naive performance. However, performance in the same-inverted condition – i.e., the condition that used familiar stimuli – was superior to average performance in the other three conditions ($t(23) = 4.47$, $p < .0001$), indicating exemplar-specificity of learning.

Overall, although there was modest transfer across conditions, the trained conditions were at a clear advantage relative to the other conditions on Day 2 for both groups, confirming the pattern shown in Figure 2.6. In this experiment, generalization to novel stimuli is consistent with the task-general learning that we found in Experiment 1. The generalization across orientation that was shown in Experiment 2, in the comparisons between old and novel faces at the untrained orientation, was absent with the phase-scrambled stimuli used here. This result suggests that familiar (or regular) stimulus structure might play a role in facilitating transfer of learning across orientation.

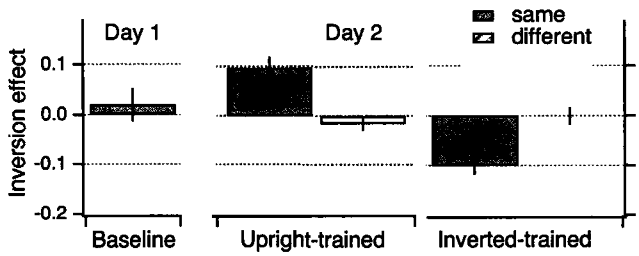


Figure 2.8: Inversion effect (upright minus inverted) for the phase-scrambled faces measured in Experiment 2. The Baseline inversion effect measured on Day 1 was calculated by comparing performance of different groups of subjects; the error bar represents the standard error of the difference between means. On Day 2 the inversion effect was calculated by comparing performance within each subject in the upright-trained and inverted-trained groups. Error bars on Day 2 represent ±1 standard error of the mean.

2.12.3 Inversion effects

One consequence of the stimulus- and orientation-specificity of learning was the emergence of an inversion effect for the trained textures (Figure 2.8). On Day 1, there was no inversion effect, because performance in the upright-trained and inverted-trained groups did not differ ($t(47) = .60, p = .55$). This result is not surprising because the stimuli did not have a canonical orientation. On Day 2, in the same-texture condition, there was a positive inversion effect for the upright-trained group ($t(23) = 4.51$, $p = .0001$), and a negative inversion effect for the inverted-trained group ($t(23) = 5.51$, $p < .0001$).

The size of the inversion effect was the same for both groups (10%). Again, this result is not surprising because the original classification of the stimuli as upright and inverted was arbitrary, and therefore we expected any induced inversion effects for the two groups to have opposite sign but approximately equal magnitude. For both groups, the inversion effect in the different-texture condition did not differ from zero (upright-trained: $t(23) = 1.06$, $p = .29$; inverted trained: $t(23) = .04, p = .96$). Hence, inversion effects on Day 2 were obtained only with textures that were seen on Day 1.

2.13 Discussion

Perceptual learning of textures was specific to the trained exemplars and orientations. Unlike what was found with the faces, there was virtually no transfer of learning across orientation except where Bin 1 was used as baseline. These results suggest that abolishing the spatial structure in faces precluded transfer of learning across orientation. Exemplar- and orientation-specificity of learning produced an inversion effect for the trained textures, consistent with what was found by Husk et al (2007) using specially derived house stimuli and band-pass limited noise stimuli. A positive inversion effect emerged after upright-training, and a negative inversion effect emerged after inverted training. The inversion effects generated despite the absence of prototypical structure in the textures argue against the proposal that such structure is essential for inversion effects (McLaren, 1997). These results indicate that canonical structure notwithstanding, familiarity with the stimulus set in whichever orientation, is the driving factor in generating an identification advantage for the given orientation.

2.14 General Discussion

The overall effects of practice on a face identification task were largely constrained to the trained exemplars and orientation, although there was some generalization of learning across orientations. This pattern of specificity suggests it is possible to fine tune the representations of individual upright and inverted faces. Greater specificity of learning was obtained with textures in which the first-order structure of normal faces had been removed. With textures, transfer across items and orientations was absent, except for when completely naive performance (Bin 1) was treated as baseline. The larger generalization across orientation found with faces relative to textures suggests that the familiar structure in faces may have facilitated some transfer across orientation. Despite the ostensibly different strategies involved in learning faces and textures, exemplar-specificity emerges as a reliable characteristic of performance improvements. Apparently, generalization of learning is limited even for expert visual tasks such as face-identification (but see Tanaka et al. (2005) for an example of generalized learning across exemplars of another expert object class using a different training paradigm). Additionally, the same amount of practice yielded equivalent changes in the size of the inversion effect both with faces and textures (i.e., after training, the inversion effect changed by 10% for all groups tested).

The extent of generalization across exemplars and orientations was influenced by the choice of baseline. When overall accuracy on Day 1 was considered baseline (Figures 3 and 6), there was no generalization across exemplars or orientation on Day 2, both with faces and with textures. However, when Bin 1 was treated as baseline, (Figures 4 and 7), there was evidence for generalization across exemplars and orientations for all groups. Overall accuracy on Day 1 includes the improvements that occurred across all 840 trials on Day 1, whereas accuracy at Bin 1 represents completely naive performance because it was based on the first 105 trials performed on the task. Relative to completely naive performance, accuracy at Bin A on Day 2 was about 10% better in the untrained conditions. However, accuracy in the trained conditions was 20-30% better than completely naive performance. We therefore interpret the generalization relative to Bin 1 as the task-general component of learning, which is obscured when overall accuracy on Day 1 is considered baseline. Evidently, generalization of learning is clearer when performance is examined at a fine resolution. Additionally, with faces, there was evidence for generalization beyond the task-general component of learning. Namely, there was generalization across orientation revealed by the advantage of old over new faces at the untrained orientation. This result was not found with textures, suggesting that transfer across orientation may have arisen due to the structural differences between faces and textures, or due to familiarity with faces more generally. In future experiments, textures with varying degrees of prototypical information (e.g., see Rousselet et al. 2008) could be used to test whether transfer across orientation increases with the amount of spatial regularity within the object

The substantial within-session learning in these tasks differs from the minimal within-session effects reported in many studies of perceptual learning. For example, negligible amounts of within-session learning, but significant between-session learning, has been found in experiments using texture discrimination (Mednick et al., 2005) and orientation discrimination (Schoups et al., 1995) tasks. However, not all studies have failed to find within-session learning effects: Poggio et al. (1992), for example, reported that accuracy in a hyperacuity task increased significantly over the course of a few hundred trials (also see Fahle et al. 1995). Hussain et al. (2008) speculated that differences in the time-course of learning could be due to differences in the experimental methods used in different experiments (e.g., method of constant stimuli in the present experiments versus method of descending limits in the texture discrimination task). However, the factors contributing to the relative magnitudes of within- and between-session perceptual learning are poorly understood. Also unknown, is whether the stimulus-specific and task-general components of learning share a similar time course in these tasks, or whether the earlier improvements are purely task-related. Future experiments may address this issue by examining stimulus-specificity of learning after varying amounts of practice.

2.14.1 Orientation-specific learning of complex stimuli

Orientation-specificity, where found in perceptual learning, has been taken as evidence for the fine-tuning of early visual mechanisms (Karni and Bertini, 1997). Studies reporting orientation-specificity typically use simple visual stimuli that vary along a single dimension (e.g., spatial frequency). For example, orientation-specific learning has been reported for discrimination of sinusoids (Fiorentini and Berardi, 1981), orientation discrimination of tilted Gabors (Schoups et al., 1995), and acuity judgements of vernier targets (Poggio et al., 1992). In such cases, it may be plausible to localize learning to cells in early cortical areas that encode the relevant properties of the stimuli. However, the physiological substrate of learning that occurs with complex patterns like faces and textures is less obvious. It is possible that orientation-specificity is a property of learning throughout the visual hierarchy, even in higher areas such as inferior temporal cortex (IT), which encode entire objects as well as their individual attributes (Logothetis et al., 1995; Desimone et al., 1984). Indeed, areas such as IT have been implicated in visual learning of complex tasks (Jagadeesh et al., 2001). We have recently reported how learning of textures, in addition to orientation, is specific to contrast polarity (Hussain et al., 2009), which is consistent with the interpretation that learning of textures engages higher visual areas.

2.14.2 Perceptual learning and configural processing

It has been proposed that expertise with a given object class enhances sensitivity to spatial configurations of features, engaging mechanisms that are not typically used for other objects (Gauthier and Tarr, 1997). Although some evidence supports this proposal in humans (Gauthier and Tarr, 1997; Tanaka et al., 2004), and in monkeys (Baker et al., 2002), other evidence does not (Robbins and McKone, 2007). In the current study, it is difficult to attribute better performance with the trained faces to increased configural processing because of the similar pattern of results found with textures. With textures, there is no common configuration of features, so it is not clear how learning could enhance configural processing for the scrambled stimuli in a manner analogous to the oft-hypothesized configural processing of faces. The alternative, which is that subjects simply get better at discriminating individual features during learning, is consistent with a report that dog experts do not show an increased reliance on configural information relative to non-experts (Robbins and McKone, 2007), and that training on inverted face identification does not yield a greater use of configural cues (Robbins and McKone, 2003). This alternative is also consistent with the results of experiments that measured classification images in face and texture discrimination tasks during the course of training (Gold et al., 2004). The classification images show that with practice, an increased stimulus area is used to discriminate faces and texture patterns. However, the increases in information-use are restricted to the local stimulus regions used initially by each subject before learning, rather than spanning across the stimulus extent. Perhaps more importantly, the relative weights of information within the local regions seem to shift with learning, so that observers become more "ideal" in their use of information. This type of local information-use might underlie the specificity of learning observed in the current experiments with upright and inverted faces, and phase-scrambled faces.

2.14.3 Perceptual learning and the face-inversion effect

The face-inversion effect has been taken as a measure of special processes engaged only by faces (Scapinello and Yarmey, 1970), namely configural encoding (McKone et al., 2007; Collishaw and Hole, 2000). Our results suggest caution with such an interpretation, as has been suggested elsewhere (Maurer et al., 2002; Sekuler et al., 2004; Husk et al., 2007; Gaspar et al., 2008a; Jiang et al., 2006; Gaspar et al., 2008b). First, we found that the decrease in the size of the inversion effect after training with inverted faces was coupled with no increase in performance with upright faces for that group. A smaller inversion effect typically would be taken to indicate less configural processing, but in this case performance with upright faces was unchanged relative to baseline. Second, as mentioned earlier, we obtained an inversion effect with textures, which have no clear configural structure. The diminished inversion effect after training with inverted faces, and the emergence of an inversion effect for the textures show that there is nothing special in the configuration of upright faces per se that elicits the inversion effect more generally.

One might still argue that the inversion effects found with faces are qualitatively different than those that arise with learning of arbitrary texture stimuli. Faces exhibit a baseline inversion effect not found with textures, or for that matter houses, prior to training. This raises the question of whether a baseline effect could be generated after training for texture patterns in which some average structure has been introduced, an issue for future studies to address.

2.14.4 Perceptual learning and norm-based coding

Another way to conceptualize improved face identification is in terms of norm-based coding (Rhodes and McLean, 1990; Rhodes and Jeffery, 2006; Rhodes et al., 2004; Burton et al., 2005). According to this framework, faces are distinguished by virtue of their position and distance from each other, or from the average face, and the average face is the accumulation of all faces in one's experience. Faces that are closely clustered resemble each other, and distance from the average enhances the distinctiveness of each face. This framework has been used to explain, for example, the other-race effect (Byatt and Rhodes, 1998). It has been suggested that perceptual learning calibrates the face-space to optimize differentiation by capturing the variance in faces that have been experienced in the subject's lifetime (Valentine, 1991). In effect, perceptual learning increases the inter-stimulus distance (and/or the distance of each face from the average), which disambiguates the faces, and improves identification. According to this scheme, inverting a face merely increases task difficulty due to larger error associated with correctly locating the face in face-space (Valentine, 1991). In this respect, inverting a face is no different than any other transformation that increases task difficulty (e.g., contrast reversal). Indeed, classification images show that similar regions of the face are used in upright and inverted face identification, but subjects are simply less efficient at extracting the information from those regions when the faces are inverted (Sekuler et al., 2004), and the level of inefficiency for inverted faces is similar to that of contrast-reversed faces (Gaspar et al., 2008a). It is now clear that inverted face identification can be made more efficient, implying that inverted face representations are refined with practice. We also show some cross-orientation transfer of learning in overall accuracy from upright to inverted faces, suggesting that upright and inverted representations are interdependent, contrary to the suggestion that upright and inverted faces are represented independently (Rhodes et al., 2004). Our experiments do not directly address the status of the average face, but the results with textures suggest that inter-item distance is sufficient to characterize performance improvements, because the textures do not conform to a clear average.

Exemplar-based models of face coding that do not incorporate the average have been discussed elsewhere (Valentine, 1991).

2.14.5 Learning of images in naturalistic conditions

Does learning in the real world differ from that found in the fairly reduced conditions used in these experiments? In naturalistic conditions, objects are typically viewed in a variety of viewpoints and illuminations, which might produce more generalization of learning across variations in the image. Furthermore, identification in naturalistic conditions could expose differences in the way images with differing spatial characteristics (such as faces and textures) are learned. These possibilities can, in fact, be explored in the lab: future experiments might test generalization of learning after observers have been exposed to several variations of a set of exemplars.

2.14.6 Conclusions

Perceptual learning of face- and texture-identification is largely orientation-specific and exemplar-specific, although faces, but not textures, show some generalization of learning across orientation. The net effect of such learning is to increase the face-inversion effect after practice with upright faces and decrease the effect after practice with inverted faces; in both cases the changes are confined to the trained face set. With textures, which have no baseline inversion effect due to the absence of a canonical orientation, the above pattern of learning yields a positive inversion effect after upright training, and a negative inversion effect after inverted training, again only for the trained stimulus set. Inversion effects appear to be driven more by familiarity than by the spatial structure of the stimuli.

References

Anderson, N. D., Wilson, H. R., 2005. The nature of synthetic face adaptation. Vision Res 45 (14), 1815–1828.

Baker, C. I., Behrmann, M., Olson, C. R., 2002. Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. Nat Neurosci 5 (11), 1210–1216.

Ball, K., Sekuler, R., 1987. Direction-specific improvement in motion discrimination. Vision Res 27 (6), 953–965.

Brainard, D. H., 1997. The psychophysics toolbox. Spat Vis 10 (4), 433–436.

Burton, A. M., Jenkins, R., Hancock, P. J. B., White, D., 2005. Robust representations for face recognition: the power of averages. Cognit Psychol 51 (3), 256–284.

Byatt, G., Rhodes, G., 1998. Recognition of own-race and other-race caricatures: implications for models of face recognition. Vision Research 38, 2455–2468.

Byatt, G., Rhodes, G., 2004. Identification of own-race and other-race faces: implications for the representation of race in face space. Psychon Bull Rev 11 (4), 735–741.

Chung, S. T. L., Levi, D. M., Li, R. W., 2006. Learning to identify contrast-defined letters in peripheral vision. Vision Res 46 (6-7), 1038–1047.

Collishaw, S., Hole, G., 2000. Featural and configurational processes in the recognition of faces of different familiarity. Perception 29, 893–909.

Crist, R., Kapadia, M., Westheimer, G., Gilbert, C., 1997. Perceptual learning of spatial localization: Specificity for orientation, position and context. Journal of Neurophysiology 78 (6), 2889–2894.

Desimone, R., Albright, T., Gross, C., Bruce, C., 1984. Stimulus-selective properties of inferior temporal neurons in the macaque. The Journal of Neuroscience 4 (8), 2051–2062.

Diamond, R., Carey, S., 1986. Why faces are and are not special: an effect of expertise. J Exp Psychol Gen 115 (2), 107–117.

Dolan, R. J., Fink, G. R., Rolls, E., Booth, M., Holmes, A., Frackowiak, R. S., Friston, K. J., 1997. How the brain learns to see objects and faces in an impoverished context. Nature 389 (6651), 596–599.

Elliott, E., Wills, E., Goldstien, A., 1973. The effects of discrimination training on the recognition of white and oriental faces. B Psychonomic Soc 2, 71–73.

Fahle, M., 2004. Perceptual learning: A case for early selection. Journal of Vision 4 (10), 879–890.

Fahle, M., Edelman, S., Poggio, T., 1995. Fast perceptual learning in hyperacuity. Vision Res 35 (21), 3003–3013.

Farah, M. J., Tanaka, J. W., Drain, H. M., 1995. What causes the face-inversion effect? J Exp Psychol Human 21 (3), 628–634.

Fiorentini, A., Berardi, N., 1981. Learning in grating waveform discrimination: specificity for orientation and spatial frequency. Vision Research 21 (7), 1149–1158.

Friere, A., Lee, K., Symons, L. A., 2000. The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. Perception 29, 159–170.

Furmanski, C. S., Engel, S. A., 2000. Perceptual learning in object recognition: object specificity and size invariance. Vision Res 40 (5), 473–484.

Gaspar, C., Bennett, P., Sekuler, A., 2008a. The effects of face inversion and contrast-reversal on efficiency and internal noise. Vision Research 48 (8), 1084–1095.

Gaspar, C., Sekuler, A., Bennett, P., Oct 2008b. Spatial frequency tuning of upright and inverted face identification. Vision Research.

Gauthier, I., Tarr, M. J., 1997. Becoming a "greeble" expert: exploring mechanisms for face recognition. Vision Res 37 (12), 1673–1682.

Gilbert, C., 1994. Early perceptual learning. Proceedings of the National Academy of Sciences of the Unites States of Americe 91, 1195–1197.

Gold, J., Bennett, P. J., Sekuler, A. B., 1999a. Identification of band-pass filtered letters and faces by human and ideal observers. Vision Res 39 (21), 3537–3560.

Gold, J., Bennett, P. J., Sekuler, A. B., 1999b. Signal but not noise changes with perceptual learning. Nature 402 (6758), 176–178.

Gold, J. M., Sekuler, A. B., Bennett, P. J., 2004. Characterizing perceptual learning with external noise. Cognitive Sci 28, 167–207.

Goldstein, A., Chance, J., 1985. Effects of training on japanese face recognition: Reduction of the other-race effect. B Psychonomic Soc 23, 211–214.

Husk, J. S., Bennett, P. J., Sekuler, A. B., 2007. Inverting houses and textures: investigating the characteristics of learned inversion effects. Vision Res 47 (27), 3350–3359.

Hussain, Z., Sekuler, A., Bennett, P., 2008. Robust perceptual learning of faces in the absence of sleep. Vision Research 48 (28), 2785–2792.

Hussain, Z., Sekuler, A., Bennett, P., 2009. Contrast reversal abolishes perceptual learning. Journal of Vision 9 (4), 1–8.

Jagadeesh, B., Chelazzi, L., Mishkin, M., Desimone, R., 2001. Learning increases stimulus salience in anterior inferior temporal cortex of the macaque. Journal of Neurophysiology 86, 290–303.

Jeffery, L., Rhodes, G., Busey, T., 2006. View-specific coding of face shape. Psychol Sci 17 (6), 501–505.

Jiang, X., Rosen, E., Zeffiro, T., VanMeter, J., Blanz, V., Riesenhuber, M., April 2006. Evaluation of a shape-based model of human face discrimination using fmri and behavioral techniques. Neuron 50, 159–172.

Karni, A., Bertini, G., 1997. Learning perceptual skills: behavioral probes into adult cortical plasticity. Curr Opin Neurobiol 7 (4), 530–535.

Leopold, D. A., O'Toole, A. J., Vetter, T., Blanz, V., 2001. Prototype-referenced shape encoding revealed by high-level aftereffects. Nat Neurosci 4 (1), 89–94.

Logothetis, N., Pauls, J., Poggio, T., 1995. Shape representation in the inferior temporal cortex of monkeys. Current Biology 5 (5), 552–563.

Maurer, D., Le Grand, R., Mondloch, C., 2002. The many faces of configural processing. Trends Cogn Sci 6 (6), 255–260.

McKone, E., Brewer, J. L., MacPherson, S., Rhodes, G., Hayward, W. G., 2007. Familiar other-race faces show normal holistic processing and are robust to perceptual stress. Perception 36 (2), 224–248.

McLaren, I. P., 1997. Categorization and perceptual learning: an analogue of the face inversion effect. Q J Exp Psychol A 50 (2), 257–273.

Mednick, S. C., Arman, A. C., Boynton, G. M., 2005. The time course and specificity of perceptual deterioration. Proc Natl Acad Sci U S A 102 (10), 3881–3885.

Mollon, J. D., Danilova, M. V., 1996. Three remarks on perceptual learning. Spat Vis 10 (1), 51–58.

Moses, Y., Ullman, S., Edelman, S., 1996. Generalization to novel images in upright and inverted faces. Perception 25 (4), 443–461.

Pelli, D. G., 1997. The videotoolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10 (4), 437–442.

Poggio, T., Fahle, M., Edelman, S., 1992. Fast perceptual learning in visual hyperacuity. Science 256 (5059), 1018–21.

Rhodes, G., Jeffery, L., 2006. Adaptive norm-based coding of facial identity. Vision Res 46 (18), 2977–2987.

Rhodes, G., Jeffery, L., Watson, T. L., Jaquet, E., Winkler, C., Clifford, C. W. G., 2004. Orientation-contingent face aftereffects and implications for face-coding mechanisms. Curr Biol 14 (23), 2119–2123.

Rhodes, G., McLean, I. G., 1990. Distinctiveness and expertise effects with homogeneous stimuli: towards a model of configural coding. Perception 19 (6), 773–794.

Robbins, R., McKone, E., 2003. Can holistic processing be learned for inverted faces? Cognition 88 (1), 79–107.

Robbins, R., McKone, E., 2007. No face-like processing for objects-of-expertise in three behavioural tasks. Cognition 103 (1), 34–79.

Rousselet, G. A., Pernet, C. R., Bennett, P. J., Sekuler, A. B., 2008. Parametric study of eeg sensitivity to phase noise during face processing. BMC Neurosci 9, 98.

Rubin, N., Nakayama, K., Shapley, R., 1997. Abrupt learning and retinal size specificity in illusory-contour perception. Curr Biol 7 (7), 461–467.

Scapinello, K., Yarmey, A., 1970. The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. Psychonomic Science 21 (6), 329–331.

Schoups, A. A., Vogels, R., Orban, G. A., 1995. Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. J Physiol 483 (Pt 3), 797–810.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., Bennett, P. J., 2004. Inversion leads to quantitative, not qualitative, changes in face processing. Curr Biol 14 (5), 391–396.

Sigman, M., Gilbert, C. D., 2000. Learning to find a shape. Nat Neurosci 3 (3), 264–269.

Tanaka, J. W., Curran, T., Sheinberg, D. L., 2005. The training and transfer of real-world perceptual expertise. Psychol Sci 16 (2), 145–151.

Tanaka, J. W., Kiefer, M., Bukach, C. M., 2004. A holistic account of the own-race effect in face recognition: evidence from a cross-cultural study. Cognition 93 (1), B1–9.

Tyler, C. W., Liu, H. C. L., McBride, B., Kontsevich, L., 1992. Bit-stealing: How to get 1786 or more gray levels for an 8-bit color monitor. In: Rogowitz, B. E. (Ed.), Proceedings of SPIE: Human vision, visual processing and digital display III. Vol. 1666. pp. 351–364.

Valentine, T., 1988. Upside-down faces: A review of the effect of inversion upon face recognition. British Journal of Psychology 79, 471–491.

Valentine, T., 1991. A unified account of the effects of distinctiveness, inversion, and race in face recognition. Q J Exp Psychol A 43 (2), 161–204.

Valentine, T., Bruce, V., 1986. The effect of race, inversion and encoding activity upon face recognition. Acta Psychol (Amst) 61 (3), 259–273.

Yi, D.-J., Olson, I. R., Chun, M. M., 2006. Shape-specific perceptual learning in a figure-ground segregation task. Vision Res 46 (6-7), 914–924.

Yin, R. K., 1969. Looking at upside-down faces. J Exp Psychol 81 (1), 141–145.

Chapter 3

Contrast-reversal abolishes perceptual learning

3.1 Abstract

We tested the effects of contrast reversal on perceptual learning in a 10AFC texture identification task. Four groups of subjects performed the task on two consecutive days. One group saw the same textures on both days, whereas three other groups saw novel, rotated (180 deg), or contrast-reversed textures on the second day. Response accuracy improved during the first day in all groups. Accuracy decreased significantly at the start of Day 2 in the groups who saw novel, rotated or contrast-reversed textures, but not in the group who saw the same textures. Moreover, the drop in performance was the same in the groups who saw novel, rotated, and contrast-reversed textures. Control experiments showed that making subjects aware of the stimulus transformations at the start of either the first or second day did not alter the results. Hence, the effects of contrast-reversal and 180 deg rotation on the generalization of learning were the same as the effect of using novel stimuli, and knowledge of the stimulus transformation did not reduce their effects. We consider the implications of this pattern of results for the neural mechanisms recruited during the identification and learning of two-dimensional visual patterns.

Citation: Hussain, Z., Sekuler, A. B., & Bennett, P. J. (2009). Contrast-reversal abolishes perceptual learning. *Journal of Vision,* 9(4):20, 1-8, http://journalofvision.org/9/4/20/

3.2 Introduction

Practice improves performance in a variety of visual tasks. With basic low-level tasks, which require simple stimuli to be discriminated about a single attribute such as orientation or spatial frequency, the effects of practice typically are specific to the trained items (Fiorentini and Berardi, 1981; Matthews et al., 1999; Ball and Sekuler, 1987; Schoups et al., 1995; Fahle and Morgan, 1996; Crist et al., 1997), and, in those situations where it has been tested, to the particular locations of the items during training (Sowden et al., 2002; Karni and Sagi, 1991; Fahle et al., 1995; Schoups et al., 1995). These behavioural effects suggest that perceptual learning may alter the response properties of early visual areas, where cells are retinotopically organized and encode basic attributes of visual stimuli (Gilbert, 1994; Karni and Bertini, 1997; Crist et al., 1997; Fahle, 2004). Indeed, physiological and neuroimaging studies reveal changes to response properties of neurons in primary visual cortex (V1) after training on low-level tasks (Pourtois et al., 2008). Similarly, the behavioural effects obtained with complex patterns offer potential insights to the visual mechanisms that encode those patterns, and are engaged during learning.

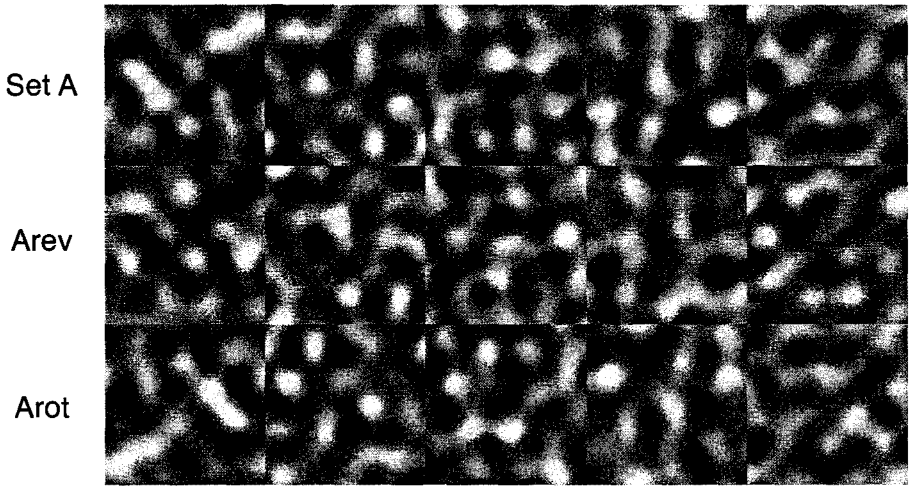


Figure 3.1: Examples of the texture stimuli. Each texture was created by applying an isotropic, band-pass (2-4 cy/image) ideal spatial frequency filter to gaussian white noise. The first row shows five of the ten textures from Set A. The second row shows the same five textures reversed in contrast. The third row shows the textures from the first row after being rotated by 180 deg.

Complex visual patterns can be discriminated on the basis of more than one attribute, such as the shape, location and orientation of features. As with simple stimuli, practice improves the discrimination and identification of complex, and even completely unfamiliar objects. For example, practice significantly improves response accuracy on a 10-AFC texture identification task (Gold et al., 1999, 2004), and the effects of practice are much greater for familiar textures (i.e., those seen during practice) than novel textures with similar spatial properties as the trained items (Hussain et al., 2005). In other words, perceptual learning in a texture identification task exhibits stimulus-specificity that is similar to the specificity reported for the aforementioned low-level tasks. With simple visual tasks, the effects of learning are clearly restricted to the single attribute of the stimulus that is being discriminated, with relatively straightforward inferences regarding the neural representation. With the identification of complex patterns such as two-dimensional textures, it is less clear what is being learned. Learning of complex patterns could involve changes at the earliest neural levels, where simple stimulus attributes are coded, or later in the visual pathway, where entire objects are represented (Desimone et al., 1984; Logothetis et al., 1995), or both. In this paper we investigate the question of what subjects learn about a given set of textures, and use the pattern of learning to consider the possible neural representations.

Response classification studies have shown that learning increases the efficiency with which subjects extract information from localized regions within the textures (Gold et al., 2004). One possibility is that subjects learn the locations and/or shapes of the most informative blobs within each texture. By examining the extent to which learning transfers across various stimulus transformations, we can gain insight into the learned representations. If the learned representation codes for shape and location, performance after learning should be invariant to image transformations such as contrast-reversal that preserve the locations and shapes of the texture blobs (see Figure 3.1, Set A vs. Arev). This manipulation leaves intact the spatial distribution of information within the image, with the contrast-defined features remaining in the same location across all images. Similarly, if the learned representation codes for global shape in an orientation-invariant manner, then we should expect performance after learning to be invariant to image transformations such as 180-deg rotations (see Figure 3.1, Set A vs. Arot). Here, we assess the effects of such image transformations on perceptual learning of texture identification.

3.3 Methods

3.3.1 Subjects

Fifty-six McMaster University undergraduate students participated in this experiment. All subjects had normal or corrected-to-normal Snellen visual acuity. The mean age and years of education were, respectively, 19.48 (SD = 2.59) and 15.37 (SD = 2.54). All subjects were compensated for their participation with a small stipend ($10/hour) or partial course credit for participating in the experiment, and all subjects were naive with respect to the aims of the experiment and had no previous experience in this task.

3.3.2 Apparatus and Stimuli

Stimuli were generated in Matlab (v. 5.2) using the Psychophysics and Video Toolboxes (Brainard, 1997; Pelli, 1997), and displayed on a 21" Sony Trinitron monitor (1024 x 768 pixels) at a frame rate of 85 Hz. Average luminance was 62.5 cd/m^2. Display luminance was measured with a PhotoResearch PR650 photometer, and the calibration data were used to build a 1779-element lookup table (Tyler et al., 1992). Customized computer software constructed the stimuli on each trial by selecting the appropriate luminance values from the calibrated lookup table and storing them in the display's eight-bit lookup table.

The textures were band-limited noise patterns created by applying an isotropic, ideal band-pass (2-4 cy/image) spatial frequency filter to white gaussian noise (Figure 3.1). Stimulus size was 256 x 256 pixels, subtending 4.8 x 4.8 deg of visual angle from the viewing distance of 114 cm. Two sets (A and B) of ten textures were created, as well as contrast-reversed (A_{rev}, B_{rev}) and 180 deg rotated (A_{rot}, B_{rot}) versions of each set. During the experiment, stimulus contrast was varied across trials using the method of constant stimuli. Seven levels of contrast were spaced equally on a logarithmic scale across a range that was sufficient to produce significant changes in performance in virtually all subjects. The textures were shown in one of three levels (low, medium and high) of static two-dimensional Gaussian noise (contrast variance = .001, .01, or .1). Hence, subjects viewed each texture at a signal-to-noise ratio that varied significantly across trials. There were 21 different stimulus conditions (seven contrast levels x three external noise levels), and these 21 conditions were randomly intermixed within a session.

3.3.3 Procedure

All subjects performed two sessions of a texture identification task at the same time on consecutive days. There were four groups, each consisting of 14 subjects. The Same group saw the same set of ten textures on both days: Seven subjects saw set A on both days, and seven subjects saw set B on both days. The Novel group transferred across sets A and B from one day to the next: From Day 1 to Day 2, seven subjects transferred from Set A to Set B, and seven subjects transferred from Set B to Set A. The Contrast-Reversed group performed Day 1 with one set of textures (i.e., A, B, A_{rev}, or B_{rev}) and Day 2 with the same set of ten textures reversed in contrast polarity. At least three subjects were assigned to each order (e.g., A-A_{rev}, B-B_{rev}, A_{rev}-A, or B_{rev}-B), with an additional two subjects randomly assigned to one of the four orders. Finally, the Rotated group performed Day 1 with one set of textures (A, B , A_{rot}, B_{rot}) and Day 2 with that same set rotated by 180 deg. At least three subjects were assigned to each order (e.g., A-A_{rot}, B-B_{rot}, A_{rot}-A, or B_{rot}-B), with an additional two subjects randomly assigned to one of the four orders. Subjects in the Novel, Contrast-Reversed, and Rotated groups were not informed that the stimuli on Day 2 differed from those seen during Day 1.

Subjects were seated in a darkened room 114 cm away from the monitor. Viewing was binocular, and viewing position and distance were stabilized with an adjustable chinrest. The experiment started after a 60 s period during which the subject adapted to the average luminance of the display. A trial began with the presentation of a black, high-contrast fixation point (0.15 x 0.15 deg) in the center of the screen for 100 ms, followed by a texture, selected randomly from one of the 21 stimulus conditions, presented for 200 ms at the center of the screen, i.e., foveally. After the texture disappeared, the entire set of 10 textures was presented as noiseless, high-contrast thumbnail images, each subtending 1.7 x 1.7 deg of visual angle. Five thumbnails were presented on the top half of the screen, and five on the bottom half, and the location of each texture in the response window was constant across trials and days. The subject's task was to inspect the thumbnail images, and decide which one of the 10 textures had been presented during the trial by clicking on the chosen texture with a computer mouse. Auditory feedback in the form of high-pitched (correct) and low-pitched (incorrect) tones informed the subject about the accuracy of each response, and the next trial began one second after presentation of the feedback. Sessions on both days comprised 40 trials per stimulus condition for a total of 840 trials. The duration of each session was 60 minutes.

3.4 Results

We calculated average proportion correct (collapsed across noise and contrast levels) on Days 1 and 2 for each of the four groups (Figure 3.2). The groups did not differ in overall accuracy during Day 1, $F(3,52) = 1.26$, $p = .3$, but there was a significant difference across groups on Day 2, $F(3,52) = 2.99$, $p = .039$. Between-session learning, defined as the difference between proportion correct on Days 1 and 2, also differed significantly across groups, $F(3,52) = 7.10$, $p = .0004$. A Tukey HSD test ($p < .05$) on between-session learning indicated that the Same group differed from all of the other groups; none of the other differences were significant. Therefore between-session learning was greatest in the Same group, and did not differ among the other three groups. This latter result suggests that contrast-reversal and rotation had the same effect on the generalization of learning as replacing the textures with a novel set.

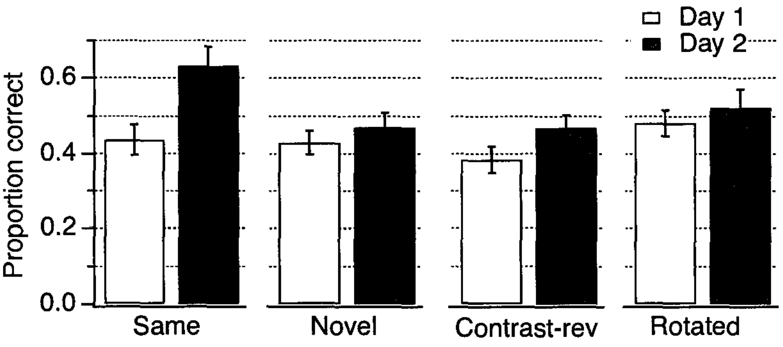


Figure 3.2: Proportion correct on Day 1 and Day 2 for each of the four groups.

The time-course of within-session learning was examined by measuring the proportion of correct responses in eight separate bins of 105 trials on Days 1 and 2 (Figure 3.3). An ANOVA on response accuracy scores measured on Day 1 yielded a significant main effect of Bin, $F(7,364) = 91.02, p < .00001$. The main effect of Group was not significant $F(3,52) = 1.25, p = .297$, nor was the Group x Bin interaction, $F(21,364) = .98, p = .48$. Within-session learning, defined as the difference between response accuracy in the first and last bins, also did not differ across groups, $F(3,52) = 1.62$, $p = .195$. However, an ANOVA on response accuracy scores measured on Day 2 did find a significant main effect of Group, $F(3,52) = 2.98$, $p = .039$, indicating that response accuracy was higher in the Same group than the other groups. As was found with the Day 1 data, the main effect of Bin was significant $F(7,364) = 61.35$, $p < .00001$, but the Group x Bin interaction was not, $F(3,364) = 1.30$, $p = .167$. Finally, within-session learning on Day 2 did not vary across groups, $F(3, 52) = .70$, $p = .55$. These analyses suggest that within-session learning was similar across groups on both days.

Figure 3.3 shows there was a drop in performance from Bin 8 to Bin 9 (across days) for three of the four groups. The drop in performance was significant for the Novel, $t(13) = -3.81$, $p = .002$, Contrast-Reversed, $t(13) = -2.81$, $p = .01$, and Rotated, $t(13) = -6.37$, $p < .0001$, groups, but not for the Same group, $t(13) = .71$, $p = .487$.

Relative to Bin 1 (i.e., initial accuracy on Day 1), performance in Bin 9 (i.e., initial accuracy on Day 2) was 8% higher for the Novel group, $t(13) = 2.98$, $p = .01$, 11% higher for the Contrast-Reversed group, $t(13) = 3.64$, $p = .002$, and 9% higher for the 180 deg rotation group, $t(13) = 2.75$, $p = .01$. For the Same group, performance at Bin 9 was 30% higher than at Bin 1, $t(13) = 7.29$, $p < .0001$. Therefore, although performance at Bin 9 was much greater for the Same group, the other groups did show some transfer-of-learning relative to completely naive performance. A one-way ANOVA on the difference scores between Bin 9 and Bin 1 indicated a significant effect of Group, $F(3, 52) = 8.39$, $p < .001$, and a Tukey HSD test indicated that the difference scores for the Same group were significantly greater than each of the other groups ($p < .01$), whereas none of the other groups differed from each other.

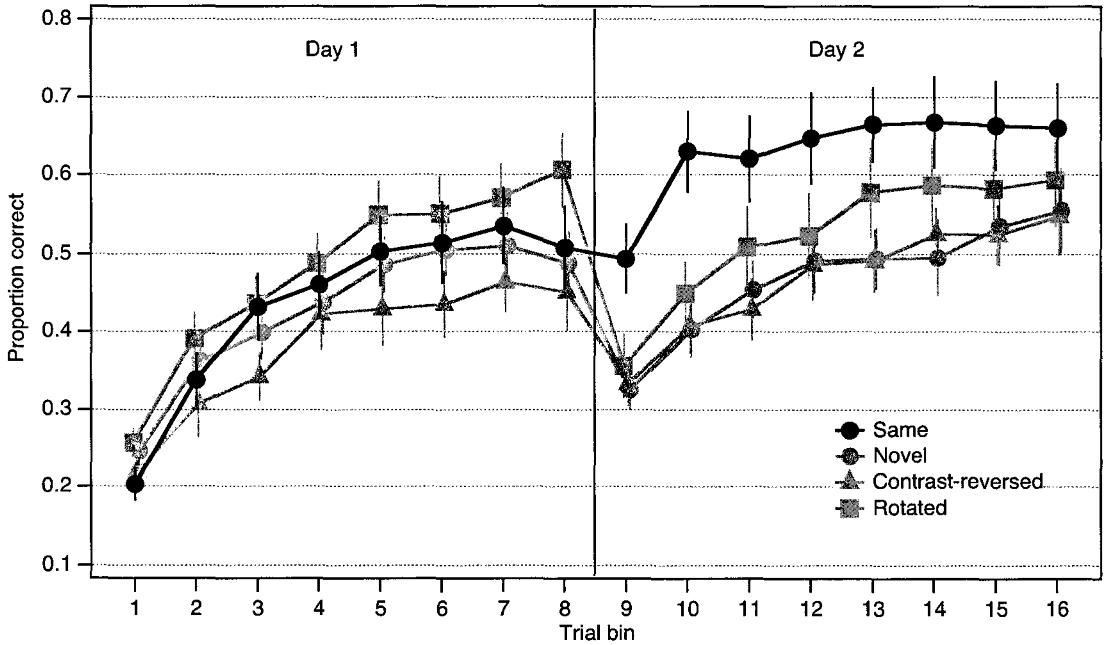
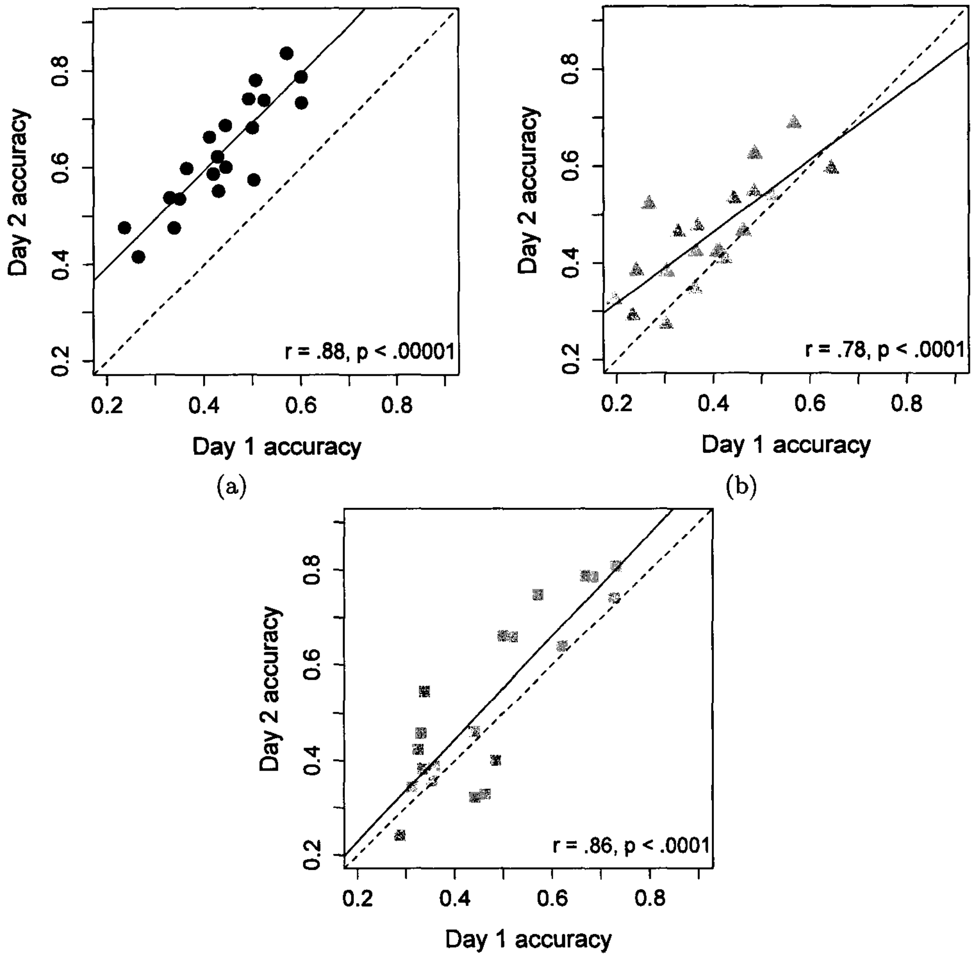


Figure 3.3: Time-course of learning on Day 1 and Day 2 for the four groups. Proportion correct is calculated at eight bins within each session, with each bin comprising 105 trials.



(a) (b) (c)

Figure 3.4: Response accuracy for individual stimuli during Days 1 and 2. a) Circles: Same textures group, b) Triangles: Contrast-Reversed group, c) Squares: Rotated group. Each point shows accuracy for a given texture, averaged across seven subjects. The dashed line in each panel represents equal performance in both sessions; the solid line is the best-fitting (least-squares) line. The Same group is the only group for which all points are above the dashed line: response accuracy for every texture increased across sessions. However, all groups show a significant positive correlation between Days 1 and 2.

For each subject, response accuracy was calculated for each texture during each session. In each group, proportion correct for each item was then averaged across subjects. Because two sets of ten items were presented to each group, this procedure yielded twenty averaged values, each based on the results from seven subjects, for each session. These values are shown in Figure 3.4 for the Contrast-Reversed, Rotated, and Same groups. In the Same group, response accuracy for individual textures was correlated strongly across sessions, $r(18) = .88$, $p < .0001$, indicating that the relative difficulty of correctly identifying individual textures was consistent across sessions. Moreover, this plot also indicates that accuracy for every texture in the Same group improved across sessions. In the Contrast-Reversed and Rotation groups, the between-session improvement was smaller and less consistent across stimuli (i.e., there was less learning). Nevertheless, response accuracy for individual items was strongly correlated in both of these groups (Contrast-Reversed: $r(18) = .77, p < .0001$; Rotated: $r(18) = .86, p < .0001$). In other words, contrast-reversal and rotation reduced learning but did not alter the relative difficulty of correctly identifying individual textures.

3.4.1 Effect of prior instructions

We examined the possibility that explicit awareness of the stimulus transformations might overcome the drop in performance from Day 1 to Day 2 found with the contrast-reversal and the rotated groups. Two control groups were tested in each condition. The Post-training Control groups were shown a printed examples of a texture stimulus and the contrast-reversed (or rotated) version of that stimulus prior to the start of the session on Day 2, after having trained with the textures on Day 1. They were told that the textures they had learned the previous day were now altered according to the shown example. Therefore, this group was aware that the textures they were seeing on Day 2 were not completely novel. The Pre-training Control groups were shown the same examples prior to training on Day 1. These subjects were told that the textures that they would see during the first session would be thus altered when they returned to perform the task the next day. Therefore, this group had the opportunity to adapt their initial learning strategy to compensate for the expected stimulus change. Care was taken to ensure that the control subjects understood the stimulus transformations. The stimuli and procedures used were the same as those described earlier. Thirty-three subjects participated in these control conditions: eight subjects in each of the two contrast reversal control groups and the rotated Post-training Control group, and nine subjects in the rotated Pre-training Control group.

Figure 3.5 shows the effect of prior instructions. We analyzed whether the instructions affected learning by separately comparing the three contrast-reversal groups (no instructions, Pre-training Control Group and Post-training Control Group), and the three rotated groups. Between-session learning (Average Day 2 - Average Day 1) did not differ across the three contrast-reversal groups, $F(2,27) = .46$, $p = .633$, nor did the difference between response accuracy in bins 8 and 9, $F(2,27) = .46$, $p = .633$. Comparisons of performance in the three rotated stimuli groups (i.e., no instructions, Pre-training Control Group and Post-training Control Group) yielded similar results: neither between-session learning, $F(2,28) = .98, p = .38$, nor the difference between bins 8 and 9, $F(2,28) = 1.626$, $p = .21$, varied significantly among the groups. Hence, there was no evidence that the effect of contrast-reversal and rotation was reduced in either of the control groups who were instructed as to the stimulus transformations on Days 1 and 2.

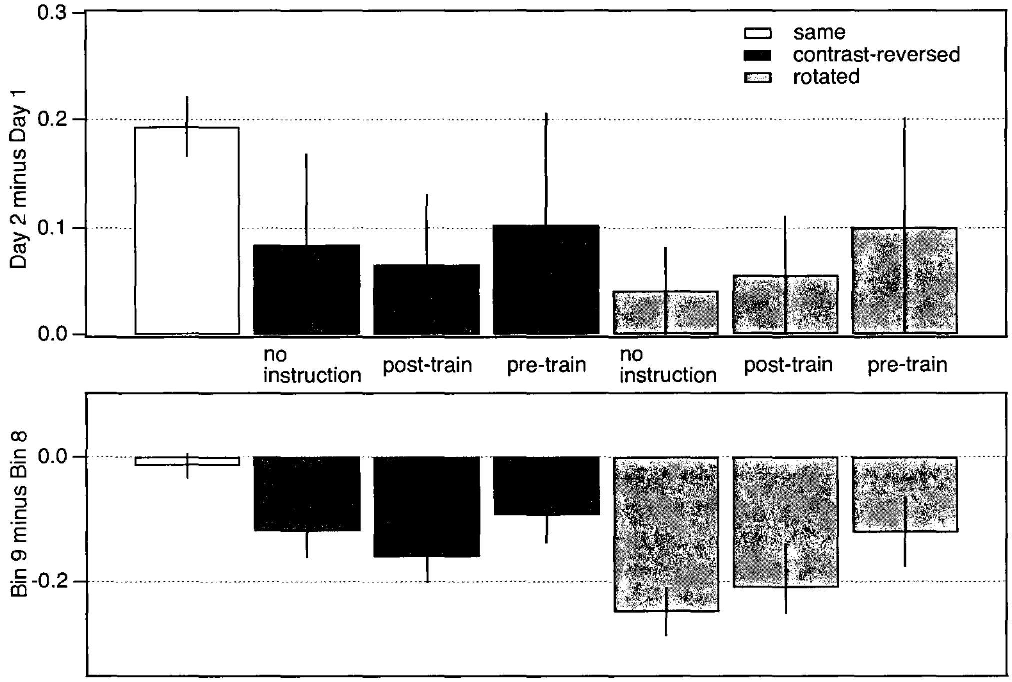


Figure 3.5: Effect of prior instructions on the drop in performance across days. Top panel: Difference between average accuracy on Day 1 and Day 2. Bottom panel: Difference between accuracy at Bin 9 and Bin 8. Data are shown for the Same group, and the three groups tested each in the Contrast-reversal and the Rotated conditions. Error bars represent ±1 standard error.

3.5 Discussion

We found that perceptual learning of texture identification did not generalize across contrast-reversal or 180 deg rotation. The large between-session learning found for the Same group (19%), was not found with the Novel, Rotation and Contrast-reversal groups; between-session learning for these three groups (approximately 5%), did not differ. Also, there was a significant, and equivalent drop in performance between Bins 8 and 9 for the Novel, Rotation, and Contrast-Reversal groups, unlike the Same group, whose performance did not decrease. These findings suggest that the effects of contrast-reversal and 180 deg rotation on the generalization of learning were similar to the effect of using a novel set of textures. In addition, the fact that performance of the control groups did not differ from that of the Contrast-Reversal and Rotation groups means that knowledge of the transformations applied to the textures did not alter their effects on generalization of learning, even when observers had an opportunity to alter their learning strategies in advance of training. Note that where performance dropped in Bin 9, it was still about 10% higher than that measured in Bin 1. That is, transforming the stimuli, or using a novel set of textures did not reduce performance to that measured initially on Day 1. We attribute the 10% savings to familiarization with the task demands, and the type of stimuli used in these experiments, with which the subjects had absolutely no experience in Bin 1. The difference between the Same group, and the other three groups in Bin 9 represents the stimulus-specific component of learning.

The effects of 180 deg rotation on the identification of similar texture patterns as the ones used here have been shown previously (Husk et al., 2007; Hussain et al., 2006). It is known from these studies that perceptual learning of textures is orientation-specific, although slight benefits do transfer to novel textures. The current experiment confirms this result, and shows that in addition to being orientation- and exemplar-specific, perceptual learning of texture patterns is sensitive to contrast polarity, suggesting that the learned representation includes information that goes beyond simply the shape and location of features, and that the nature of the learned representation is not altered by the introduction of prior knowledge about potential image transformations.

From previous work we know that with learning, the templates used in texture identification gradually become more ideal, i.e., observers use more of the available information (Gold et al., 2004), which can also be modeled as better extraction of the relevant signal (Gold et al., 1999). We now know that the optimization of templates due to learning is precise to the exposed version of the items, even when the location of differentiating features is unchanged within the stimulus set. An ideal observer would attain equal sensitivity for a given set of textures and its contrast-reversed version because the relative discriminability of items within the set is unchanged after contrast reversal. Although human performance does confirm that the relative discriminability of the stimuli is intact after rotation and contrast reversal (Figure 4), the transformed items must nevertheless be re-learned.

The current results are inconsistent with recent claims that the detrimental effects of contrast-reversal on identification are unique to faces (Nederhouser et al., 2007). Based on experiments comparing perceptual matching of faces with perceptual matching of shaded three-dimensional objects that were designed to have the same surface properties as faces, Nederhouser et al. suggested that "the representation that mediates the recognition of faces, unlike those for any other class of objects, is uniquely sensitive to contrast polarity. Human recognition of non-face objects is not sensitive to changes in contrast polarity..." (p. 2141). Our results are inconsistent with this claim, and show that the effects of contrast-reversal can be obtained with two-dimensional patterns that share neither the surface properties of faces nor the within-object class structural homogeneity of faces. Thus, neither the presence of contrast-reversal effects nor inversion effects are behavioural markers of face-specific perceptual processes.

Contrast-reversal and stimulus rotation have also been shown to disrupt the learning of texture segregation, where the task required subjects to detect the presence of a single counterphase gabor element embedded within a grid of 16 gabor elements (Grieco et al., 2006). In the segregation task used by Grieco et al. (2006), the contrast-polarity of the target was the only discriminating feature, making the task a relatively-low level task. Grieco et al. (2006) showed that perceptual learning of the segregation task was specific to the contrast polarity and orientation of the target and background, leading the authors to infer that the neural locus of learning to be odd-symmetric simple cells early in the visual pathway. Such an inference is less straightforward with respect to the present data, because the task used here was a high-level task involving complex patterns that can be identified on the basis of multiple attributes. The representation of complex patterns is thought to occur later in the visual pathway, beyond area V1, where cells are selective for the *entire object* in addition to single attributes of the object such as orientation (Logothetis et al., 1995; Desimone et al., 1984; Tanaka et al., 1991; Sary et al., 1995). In monkeys, cells from higher areas such as inferotemporal cortex (IT) are recruited during learning of unfamiliar two-dimensional patterns, and the responses of these cells are view-dependent, but invariant to changes in scale or location (Logothetis et al., 1995). The properties of these cells accord with the object-specificity and scale-invariance of object learning shown in humans (Furmanski and Engel, 2000), and with the item-specificity of texture learning shown here and elsewhere (Hussain et al., 2005). Additionally, studies with behaving monkeys have shown that although cells sensitive to contrast polarity are present in V1, the proportion of such cells is much larger beyond area V1 (Zhou et al., 2000; Ito et al., 1994). At least one study has explicitly suggested the involvement of inferotemporal cortex (IT) in coding the contrast polarity of complex patterns (Ito et al., 1994). Therefore, although the present results are consistent with those of Grieco et al (2006) in implicating even- and odd-symmetric simple cells in texture learning, we differ in suggesting that learning of this task could just as well be mediated by neurons later in visual processing, in areas such as IT.

3.6 Conclusion

Stimulus-specific effects of perceptual learning, when found with simple visual stimuli, are typically interpreted as indicating the involvement of early visual areas in tasks that require subjects to discriminate stimuli that vary along a single attribute such as orientation, spatial frequency, or direction of motion (Gilbert, 1994; Karni and Bertini, 1997; Crist et al., 1997; Fahle, 2004). Here, we show two types of specificity in learning of complex patterns: orientation specificity and contrast polarity specificity, and discuss how, aside from ostensible changes in early visual cortex such effects could arise from learning in later visual areas, such as inferotemporal cortex. The principle of stimulus specificity in learning clearly manifests itself across a range of stimulus complexities, and the specificity of coding serves as a constraint even for stimuli as complex as two-dimensional textures.

References

Ball, K., Sekuler, R., 1987. Direction-specific improvement in motion discrimination. Vision Res 27 (6), 953–965.

Brainard, D. H., 1997. The psychophysics toolbox. Spat Vis 10 (4), 433–436.

Crist, R., Kapadia, M., Westheimer, G., Gilbert, C., 1997. Perceptual learning of spatial localization: Specificity for orientation, position and context. Journal of Neurophysiology 78 (6), 2889–2894.

Desimone, R., Albright, T., Gross, C., Bruce, C., 1984. Stimulus-selective properties of inferior temporal neurons in the macaque. The Journal of Neuroscience 4 (8), 2051–2062.

Fahle, M., 2004. Perceptual learning: A case for early selection. Journal of Vision 4 (10), 879–890.

Fahle, M., Edelman, S., Poggio, T., 1995. Fast perceptual learning in hyperacuity. Vision Res 35 (21), 3003–3013.

Fahle, M., Morgan, M., 1996. No transfer of perceptual learning between similar stimuli in the same retinal position. Current Biology 6 (3), 292–297.

Fiorentini, A., Berardi, N., 1981. Learning in grating waveform discrimination: specificity for orientation and spatial frequency. Vision Research 21 (7), 1149–1158.

Furmanski, C. S., Engel, S. A., 2000. Perceptual learning in object recognition: object specificity and size invariance. Vision Res 40 (5), 473–484.

Gilbert, C., 1994. Early perceptual learning. Proceedings of the National Academy of Sciences of the Unites States of Americe 91, 1195–1197.

Gold, J., Bennett, P. J., Sekuler, A. B., 1999. Signal but not noise changes with perceptual learning. Nature 402 (6758), 176–178.

Gold, J. M., Sekuler, A. B., Bennett, P. J., 2004. Characterizing perceptual learning with external noise. Cognitive Sci 28, 167–207.

Grieco, A., Casco, C., Roncato, S., 2006. Texture segregation on the basis of contrast polarity of odd-symmetric filters. Vision Research 46, 3526–3536.

Husk, J. S., Bennett, P. J., Sekuler, A. B., 2007. Inverting houses and textures: investigating the characteristics of learned inversion effects. Vision Res 47 (27), 3350–3359.

Hussain, Z., Bennett, P., Sekuler, A., 2005. Perceptual learning of faces and textures is tuned to trained identities. Perception: ECVP Abstract Supplement 34.

Hussain, Z., Bennett, P. J., Sekuler, A. B., 2006. Face-inversion effects flex with perceptual learning. J Vision 6 (6), 153a.

Ito, M., Fujita, I., Tamura, H., Tanaka, K., 1994. Processing on contrast polarity of visual images in inferotemporal cortex of macaque monkey. Cerebral Cortex 5, 499–508.

Karni, A., Bertini, G., 1997. Learning perceptual skills: behavioral probes into adult cortical plasticity. Curr Opin Neurobiol 7 (4), 530–535.

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 of the Unites States of Americe 88 (11), 4966–4970.

Logothetis, N., Pauls, J., Poggio, T., 1995. Shape representation in the inferior temporal cortex of monkeys. Current Biology 5 (5), 552–563.

Matthews, N., Liu, Z., Geesaman, B., Qian, N., 1999. Perceptual learning on orientation and direction discrimination. Vision Research 30 (22), 3692–3701.

Nederhouser, M., Yue, X., Mangini, M., Biederman, I., 2007. The deleterious effect of contrast reversal on recognition is unique to faces, not objects. Vision Research 47, 2134–2142.

Pelli, D. G., 1997. The videotoolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10 (4), 437–442.

Pourtois, G., Rauss, K., Vuilleumier, P., Schwarz, S., 2008. Effects of perceptual learning on primary visual cortex activity in humans. Vision Research 48 (1), 55–62.

Sary, G., Vogels, R., Kovacs, G., Orban, G., 1995. Responses of monkey inferior temporal neurons to luminance-, motion-, and texture-defined gratings. Journal of Neurophysiology 73 (4), 1341–1353.

Schoups, A. A., Vogels, R., Orban, G. A., 1995. Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. J Physiol 483 (Pt 3), 797–810.

Sowden, P., Rose, D., Davies, I., 2002. Perceptual learning of luminance contrast detection: specific for spatial frequency and retinal location but not orientation. Vision Research 42 (10), 1249–1258.

Tanaka, K., Saito, H., Fukada, Y., Moriya, M., 1991. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. Journal of Neurophysiology 66 (1), 170–189.

Tyler, C. W., Liu, H. C. L., McBride, B., Kontsevich, L., 1992. Bit-stealing: How to get 1786 or more gray levels for an 8-bit color monitor. In: Rogowitz, B. E. (Ed.), Proceedings of SPIE: Human vision, visual processing and digital display III. Vol. 1666. pp. 351–364.

Zhou, H., Friedman, H., von der Heydt, R., 2000. Coding of border ownership in monkey visual cortex. The Journal of Neuroscience 20 (17), 6594–6611.

Chapter 4

Robust perceptual learning of faces in the absence of sleep

4.1 Abstract

This study examines the effects of sleep on learning in a face identification task. Five groups of subjects performed a 1-of-10 face identification task in two sessions separated by 3, 12, and 24 hours. Session 1 consisted of four blocks of 105 trials each; Session 2 consisted of eight blocks of trials. All groups exhibited significant improvement in response accuracy within each session. Furthermore, between-session learning – defined as the difference in proportion correct between sessions 1 and 2 – was significant for all groups. Between-session learning was greater in groups that slept between sessions, but the effect was small and affected performance only in the first block of trials in Session 2. Overall, we find that sleep's contribution is a small proportion of the total amount learned in face identification, with improvements continuing to accrue in its absence.

Citation: Hussain, Z., Sekuler, A. B., & Bennett, P. J. (2008). Robust perceptual learning of faces in the absence of sleep. *Vision Research*, 48(28), 2785-2792.

4.2 Introduction

Perceptual and motor skills typically improve with practice over a time-course that consists of at least two components: a fast component in which performance improves during a practice session (Fiorentini and Berardi, 1981; Poggio et al., 1992; Fahle et al., 1995; Fahle, 1994; Karni et al., 1998; Karni and Sagi, 1993), and a slow component in which the benefits of practice accumulate across sessions (McKee and Westheimer, 1978; Brashers-Krug et al., 1996; Fahle, 1994; Karni et al., 1998; Karni and Sagi, 1993; Press et al., 2005). This second, slow component, sometimes referred to as latent, or offline, learning, is affected by sleep (Atienza et al., 2004; Fenn et al., 2003; Fischer et al., 2002; Gottselig et al., 2004; Karni et al., 1994; Stickgold et al., 2000a,b; Walker et al., 2003). Results from Siegel (2001) and Walker and Stickgold (2004), for example, suggest that the REM and SWS stages of sleep are important for consolidating newly-acquired information into long-term memory (but see Song et al. 2007; Vertes and Eastman 2000; Vertes 2004). Sleep is thought to be important for establishing the stimulus-specific benefits of perceptual learning (Karni and Bertini, 1997; Karni et al., 1998; Karni and Sagi, 1993), rather than general benefits that presumably can transfer across experimental conditions (but see Fenn et al. 2003). Sleep's distinct contribution to latent learning is supported by a study of motor learning, which found that the substantial improvement in performance that is observed across days of testing – i.e., after a night's sleep – is uncorrelated with within-session effects (Walker et al., 2003). However, sleep is not necessary for latent learning, which can occur in the absence of sleep if the gap between testing sessions exceeds 4-6 hours (Fischer et al., 2002; Gottselig et al., 2004; Robertson et al., 2004; Roth et al., 2005).

In the visual domain, most demonstrations of sleep-related benefits have used a particular kind of texture discrimination task. For example, Karni and Sagi (1991) measured thresholds for discriminating peripherally-viewed horizontal and vertical textured targets, which were embedded in an array of short line segments, by varying the SOA between the offset of the target array and the onset of a mask. Using this task, Karni and Sagi (1993) found that sensitivity in a texture discrimination task increased only after at least eight hours elapsed after the initial training session. Subsequently, it was shown that the effects of practice were abolished by the interruption of sleep, with both REM and SWS contributing to the overall amount learned (Karni et al., 1994; Stickgold et al., 2000a,b). Finally, a nap taken in-between successive practice sessions reverses the deterioration that otherwise occurs with repeated testing during the same day (Mednick et al., 2003, 2002, 2005). Using a different task, Schoups et al. (1995) reported that practice lowered orientation discrimination thresholds measured with circular noise fields rotated about the oblique axis, but only when subjects slept between sessions. Learning in both the texture discrimination task and the orientation discrimination task was specific to the trained location in the visual field and, in the orientation discrimination task, to the orientations used during training. For both tasks, within-session gains were only observed early within the first session, after which performance saturated. Further improvements in performance only occurred across sessions, i.e., the improvements were latent.

The role of sleep in consolidation is yet to be demonstrated for the remaining variety of visual tasks amenable to learning. We have previously shown that performance on a one-of-ten forced choice face-identification task improves substantially across days (Gold et al., 1999b), and that the improvements are specific to the trained face exemplars and orientations (Hussain et al., 2005). Here we assess the contribution and sleep-dependence of latent learning to the overall improvement found with this task, and examine whether the amount of latent learning is associated with within-session learning.

4.3 Methods

4.3.1 Subjects

One hundred and three McMaster University undergraduate students participated in the experiment. All subjects had normal or corrected-to-normal Snellen visual acuity. The mean age and years of education were respectively 20.9 (SD = 3.31) and 16.9 (SD = 2.30). All subjects received a small fee ($10/hour) or partial course credit for participating in the experiment, and were naive with respect to the task. All subjects provided informed consent prior to the start of the experiment. Seventy-four of the subjects were female; the remaining 29 were male.

4.3.2 Apparatus and Stimuli

Stimuli were generated in Matlab (The Mathworks, v. 5.2) using the Psychophysics and Video Toolboxes (Brainard, 1997; Pelli, 1997). Stimuli were displayed on a 21" Sony Trinitron monitor at a resolution of 1024x768 pixels, which at the viewing distance of 114 cm translated to 26.7 pixels per cm. Frame rate was 85 Hz. Average luminance was 62.51 cd/m^2. The monitor calibration data were used to build a 1779-element lookup table (Tyler et al., 1992) and customized computer software constructed the stimuli on each trial by selecting the appropriate luminance values from the calibrated lookup table and storing them in the display's eight-bit lookup table.

The stimuli were faces of five males and five female faces cropped to show only internal features and equated for spatial frequency content. The methods used to create the stimuli have been described previously (Gold et al., 1999a,b). Stimulus size was 256 x 256 pixels, subtending 4.8 x 4.8 degrees of visual angle from the viewing distance of 114 cm. During the experiment, stimulus contrast was varied across trials using the method of constant stimuli. The seven levels of contrast were spaced equally on a logarithmic scale, and spanned a range that was sufficient to produce significant changes in performance in virtually all subjects. The images were shown in three levels (low, medium and high) of static two-dimensional Gaussian noise, created by sampling from distributions with variances of .001, .01, and .1. Hence, there were a total of 21 stimulus conditions (seven contrast levels x three external noise levels) that allowed subjects to view each face at a variety of signal-to-noise ratios.

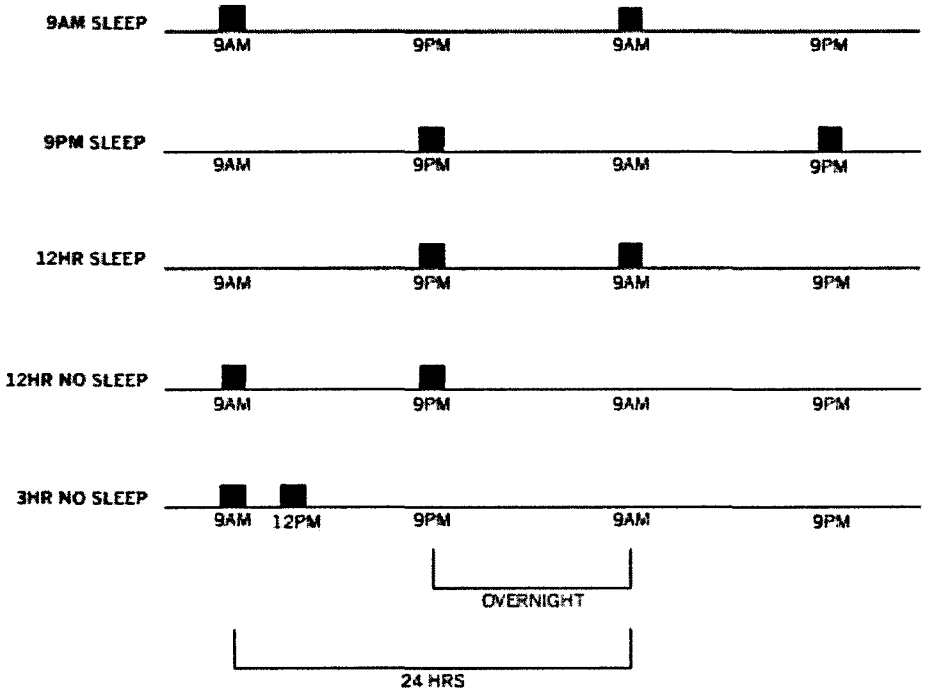


Figure 4.1: Illustration of the experimental design. Placeholders indicate the time of testing for each group across two consecutive days.

4.3.3 Procedure

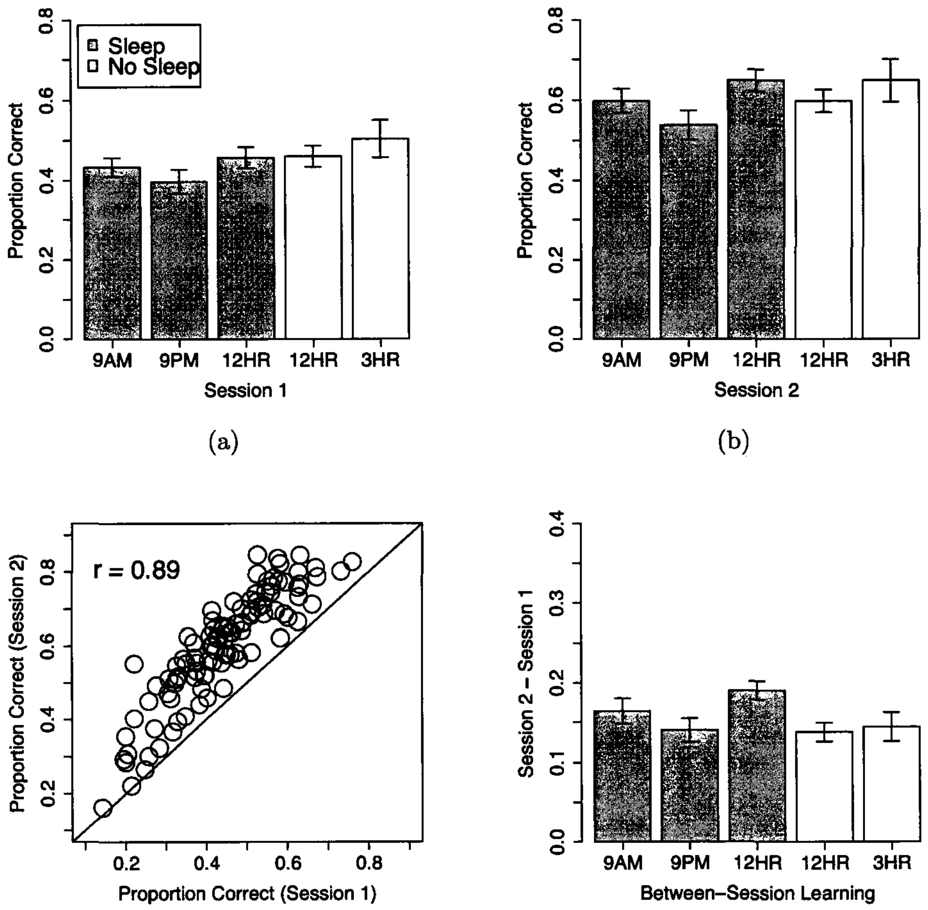
All subjects performed a face identification task in two sessions (Figure 4.1). Subjects in the 12HR-Sleep group (n = 24) performed Session 1 at 9pm and Session 2 at 9am the next day. Subjects in the 12HR-No-Sleep group (n = 24) performed Session 1 at 9am and Session 2 at 9pm on the same day. Thus, there was a 12-hour interval between sessions in both the 12HR-Sleep and 12HR-No-Sleep groups. In the 9AM-Sleep group (n = 24), subjects performed Sessions 1 and 2 at 9am on Days 1 and 2. In the 9PM-Sleep group (n = 24), subjects performed Sessions 1 and 2 at 9pm on Days 1 and 2. Therefore, there was a 24-hour interval between sessions in both the 9AM-Sleep and 9PM-Sleep groups. Subjects in the sleep groups were instructed to sleep normally overnight (between sessions), whereas subjects in the 12HR-No-Sleep group were instructed to not sleep or nap between sessions. An additional group of seven subjects in the 3HR-No-Sleep was tested with a three-hour interval between sessions (Session 1 at 9am and Session 2 at noon on the same day), with instructions not to nap between sessions.

During each session, subjects were seated in a darkened room 114 cm away from the monitor. Viewing was binocular, and viewing position and distance were stabilized with an adjustable chin/forehead rest. The experiment started after a 60 s period during which the subject adapted to the average luminance of the display. A trial began with the presentation of a fixation point – a black high-contrast spot (0.15 x 0.15 deg) – in the center of the screen for 100 ms, followed by a randomly selected face presented for approximately 200 ms at the center of the screen. After the face disappeared, the entire set of 10 faces was presented as two rows of five noiseless, high-contrast thumbnail images, each subtending approximately 1.7 x 1.7 deg. The subject's task was to decide which one of the 10 faces had been presented during the trial by selecting one of the thumbnail images with a computer mouse. The location of each face in the response window was constant across subjects, trials, and sessions. Auditory feedback in the form of high-pitched (correct) and low-pitched (incorrect) tones informed the subject about the accuracy of each response, and the next trial began one second after presentation of the feedback. Session 1 comprised 20 trials per stimulus condition for a total of 420 trials. Session 2 comprised 40 trials per stimulus condition for a total of 840 trials. Sessions 1 and 2 lasted approximately 30 and 60 minutes, respectively.

Subjects completed the Morningness-Eveningness questionnaire (Horne and Ostberg, 1976), which assesses the time of day at which the subjects perform optimally. The questionnaire was completed prior to the start of the experiment in Session 1.

4.4 Results

Statistical analyses were done with R (R Development Core Team, 2007). All *t* tests were two-tailed and assumed unequal group variances. Where appropriate, we report 95% confidence intervals (95%CI) for estimated parameters and Cohen's *f* as a measure of effect size (Cohen, 1988).



(a) (b)

(c) (d)

Figure 4.2: a) Proportion correct during Session 1. Performance did not vary significantly across groups. b) Proportion correct during Session 2. Performance did not vary significantly across groups. c) A scatter plot shown the association between proportion correct in Sessions 1 and 2. The diagonal line indicates equal performance in both sessions. Note that all points are above the line – indicating that all subjects showed some improvement across sessions. The correlation between sessions was 0.89. d) The amount learned, defined as the difference between response accuracy in Sessions 1 and 2, for each group. There was a significant interaction between Group and the amount learned.

Due to experimenter error, seven subjects did not complete the Morning-Eveningness questionnaire. Analysis of the completed questionnaires showed that subjects were unbiased as to their preferred time of day ($M = 46$, $SD = 9.18$), and that the preferred time of day did not vary across the five groups, $F(5, 90) = .8275, p = .53$. Preliminary analyses also indicated that questionnaire scores were not associated with task performance, and therefore we did not include them in subsequent analyses.

Figures 4.2a-b show proportion correct (collapsed across stimulus contrast and noise levels) in each session. In Session 1, the average proportion correct was 0.44 and did not vary significantly across groups, $f = 0.11$, $F(4, 98) = 1.32$, $p = .267$, indicating that the time of day did not affect initial performance. In Session 2, the average proportion correct was 0.59, and also did not vary significantly across groups, $f = 0.18$, $F(4, 98) = 1.85$, $p = 0.13$. Response accuracy in Session 1 was highly correlated with accuracy in Session 2, $r = 0.89$, $t(101) = 19.82$, $p < 0.0001$, demonstrating that our measures were reliable (Figure 4.2c).

Between-session learning – defined as the difference between response accuracy in Sessions 1 and 2 – is plotted in Figure 4.2d for each group. The increase in response accuracy across sessions was significant, 95%CI=(0.144, 0.170), $t(102) = 23.2$, $p < .0001$) and was highly reliable across subjects. Indeed, every subject showed an increase in response accuracy across sessions (see Figure 4.2c). There was a significant effect of Group, $f = 0.25$, $F(4, 98) = 2.55$, $p = .044$, indicating that some groups improved more than others. t tests were used to conduct orthogonal comparisons that evaluated the effects of sleep, interval between sessions, and time of day on the amount learned. A t-test that compared the groups that slept (9AM, 9PM, 12HR) to the groups that did not sleep (12HR, 3HR) found a marginally significant effect of sleep: the amount of learning was slightly greater in the sleep groups than in the no-sleep groups, 95%CI=(-0.002, 0.052), $t(101) = 1.97$, $p = .052$. A second t-test found that more learning occurred in the 12HR-Sleep group than in the two 24 hour sleep groups, 95%CI=(0.006, 0.071), $t(70) = 2.375$, $p = .021$. A third t-test that compared the 9AM-Sleep group to the 9PM-Sleep group found no effect of time-of-day on the amount learned, 95%CI=(-0.02, 0.058), $t(46) = 1.093$, $p = .28$. These analyses suggest that learning was slightly greater in subjects that slept between experimental sessions, and that the benefits of sleep were greater when the sessions were separated by twelve hours instead of twenty-four hours.

The time-course of within-session learning was examined by measuring the proportion

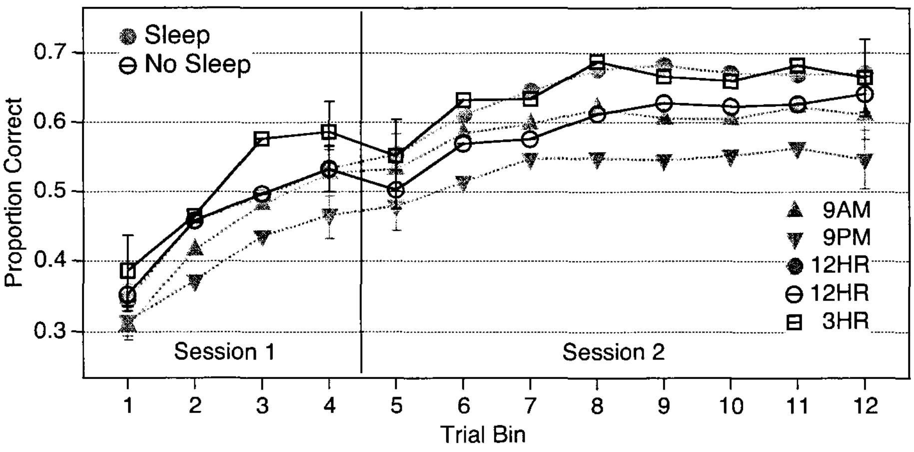


Figure 4.3: Time-course of learning across both sessions for all groups. The proportion of correct responses, averaged across subjects, is plotted for bins of 105 trials. Bins 1-4 are from Session 1, whereas bins 5-12 are from Session 2. The filled and unfilled symbols represent data from Sleep and No Sleep groups, respectively. Error bars represent ±1 standard error. For clarity, error bars are shown for four bins only. Error bars in other bins were similar to the ones shown in the figure.

of correct responses that occurred in separate bins of 105 trials. This procedure yielded four bins of trials within Session 1 and eight bins within Session 2. Figure 4.3 shows proportion correct for all groups plotted as a function of bin number. During Session 1, performance improved on average by 0.18 from bin 1 to bin 4. An analysis of variance revealed found a significant main effect of Bin, $f = 1.02$, $F(3,294) = 143.43$, $p < .0001$, but the main effect of Group, $f = 0.05$, $F(4,98) = 1.32$, $p = 0.27$, and the Group x Bin interaction, $f = 0.04$, $F(12,294) = 1.05$, $p = 0.40$, were not significant. Further analyses showed that the linear, $F(1,98) = 256.98, p < .0001$, and quadratic, $F(1,98) = 22.90, p < .0001$), trends across bins were both significant, and that neither trend interacted with Group (Group x Linear Trend: $F(4,98) = 1.02, p = 0.4$; Group x Quadratic Trend: $F(4,98) = 0.53, p = 0.71$). During Session 2, proportion correct increased by 0.1 from bin 5 to bin 12. An analysis of variance revealed a significant main effect of Bin, $f = 0.54$, $F(7,686) = 36.1$, $p < .0001$, but the main effect of Group, $f = 0.06$, $F(4,98) = 1.85$, $p = 0.13$, and the Bin x Group interaction, $f = 0$, $F(28,686) = 0.89$, $p = 0.62$, were not significant. Furthermore, the linear, $F(1,98) = 78.5, p < .0001$, and quadratic, $F(1,98) = 52.4, p < .0001$, trends were both significant, and neither trend differed significantly across groups (Group x Linear Trend: $F(4,98) = 1.48, p = 0.21$; Group x Quadratic Trend: $F(4,98) = 0.74, p = 0.57$). We also conducted a more focussed test of the effect of sleep on Day 2 by combining all subjects in the three sleep groups into one group, and all subjects in the two no-sleep groups into another group. An ANOVA on these new groups (sleep vs. no-sleep) found the same, significant main effect of Bin that was found in the previous ANOVA, a non-significant main effect of Sleep, $f = 0$, $F(1, 101) = 0.2$, $p = 0.65$, and a non-significant Sleep x Bin interaction, $f = 0$, $F(28, 686) = 0.89$, $p = 0.62$. Hence, the overall trends were similar in Sessions 1 and 2: performance improved significantly within each session, the time-course of within-session learning did not vary across groups, and there was no indication that performance differed significantly between sleep and no-sleep groups.

A close examination of Figure 4.3 suggests that the presence or absence of sleep affected performance at the start of Session 2. Specifically, performance in the No Sleep groups was less accurate in bin 5 than bin 4 (bin5 - bin4 = -0.03), and then increased in bin 6 (bin6 - bin5 = 0.07). In the sleep groups, on the other hand, performance increased monotonically across bins 4 through 6 (bin5 - bin4 = 0.01; bin6 - bin5 = 0.05). A trend analysis confirmed that the quadratic trend in performance across bins 4-6 differed significantly across groups, $F(1, 98) = 5.52, p = 0.02$. The origin of this effect is shown in Figure 4.4, which plots difference scores calculated for response accuracy in bins 4 and 5. The boxplots indicate that response accuracy was not higher in bin 5 than bin 4 in a substantial proportion of subjects in each group, and that the median difference scores were close to zero for all groups. However, the difference scores were slightly lower in the No Sleep groups. Hence, the data suggest that sleep, rather than boosting performance, prevented a slight deterioration in performance at the start of Session 2.

A global measure of within-session learning was defined as the difference between proportion correct measured in the first and last bins: for example, learning during Session 1 was the difference between proportion correct in bins 4 and 1. Surprisingly, estimates of within-session learning from Sessions 1 and 2 were not correlated, but each measure of within-session learning was correlated with between-session learning (Figure 4.5). To determine if the association of within- and between-session learning varied across groups, we evaluated the interaction term in linear models of the form

$$b = w + G + wG \tag{4.1}$$

where b is between-session learning (see Figure 4.2d), w is within-session learning, G is a factor representing Group, and wG is the $w \times G$ interaction. The analyses indicated that the association between learning during Session 1 and between-session learning did not vary across groups, $f = 0$, $F(4, 93) = 0.62$, $p = 0.69$. Likewise, the association between

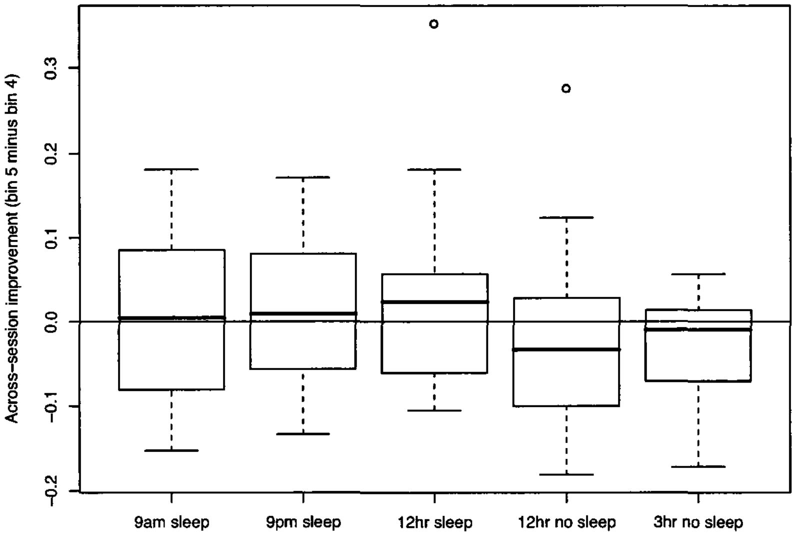
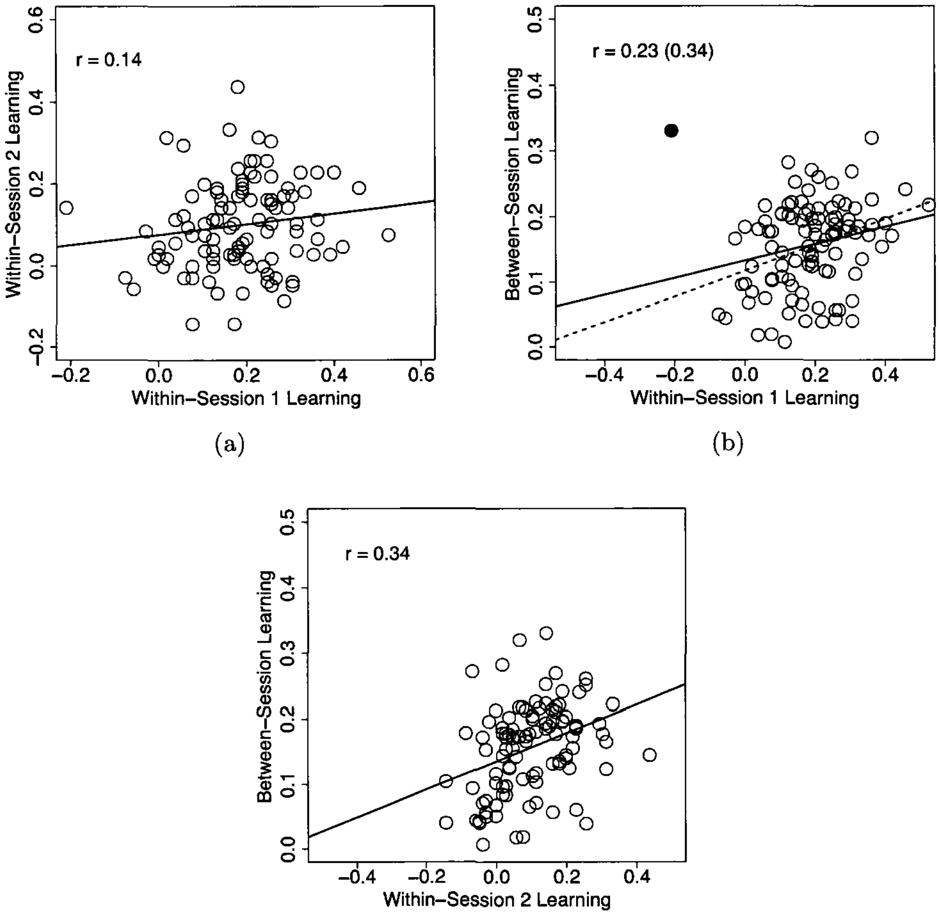


Figure 4.4: Boxplots illustrating the changes in response accuracy (i.e., bin 5 minus bin 4) for five groups. In each boxplot, the box depicts the interquartile range (IQR), the horizontal line indicates the group median, and the whiskers extend to the most extreme data points that are within 1.5 × IQR. The unfilled symbols represent outliers.



(a) (b)

(c)

Figure 4.5: Scatter plots showing the association between learning in Sessions 1 and 2. Within-session learning was defined as the difference between proportion correct in the first and last bins of a single session. Between-session learning was defined as the difference in the proportion correct between Sessions 1 and 2. Solid lines in each panel represent the best-fitting (least-squares) fit to the data. a) Within-session learning from the two sessions was not correlated ($r = 0.14, t(101) = 1.42, p = 0.16$). b) Within-session learning from Session 1 was significantly correlated with between-session learning ($r = 0.23, t(101) = 2.35, p = 0.02$). One subject, indicated by the solid symbol, was an outlier: excluding that subject increased the correlation to 0.34, $t(100) = 3.61, p = 0.0005$. The dashed line is the regression line computed after excluding the outlier. c) Within-session learning from Session 2 was significantly correlated with between-session learning ($r = 0.34, t(101) = 3.71, p = 0.0003$).

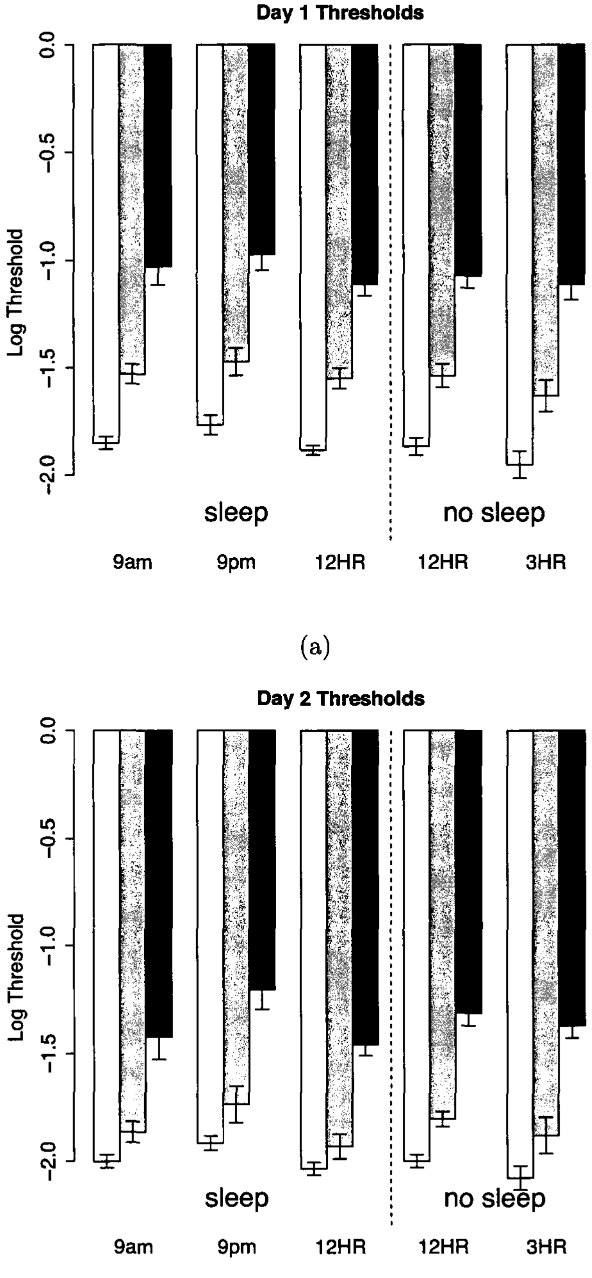
learning during Session 2 and between-session learning did not differ across groups, $f = 0.04$, $F(4, 93) = 1.40$, $p = 0.24$. Hence, the analyses show that the magnitudes of within- and between-session learning were correlated, and that correlation was similar across groups.

Hauptmann et al. (2005) reported that between-session learning in a letter enumeration task was larger in subjects whose performance had reached an asymptotic value during the first session. To determine whether a similar effect held in the current experiment, we first examined how proportion correct varied across bins 1-4 in individual subjects. We identified 45 subjects whose proportion correct was approximately constant or declined in bins 2-4; response accuracy in the remaining 58 subjects increased in bins 2-4 and showed no signs of reaching an upper asymptote. These two groups of subjects were represented by a binary classification factor that indicated if a subject did or did not reach an upper limit during Session 1. Finally, this classification factor was added to a model that predicted between-session learning from within-session learning during Session 1, Group membership, and the interaction between those two predictor variables (Equation 4.1). Adding the binary classification factor did not improve the model's fit significantly, $f = 0$, $F(1, 92) = 0.38$, $p = 0.54$. Hence, there was no evidence that the predictability of between-session learning from learning during Session 1 differed between subjects whose performance did and did not reach an upper asymptote during the first session.

4.4.1 Face Identification Thresholds

Our experiment used a fixed set of contrasts for all subjects, and therefore was not optimized to measure thresholds accurately in individuals. Nevertheless, it was possible to fit psychometric functions to all but eight of the 618 individual sets of data. The resulting thresholds, defined as the rms contrast that corresponded to 50% correct responses, exhibited significant positive skew and contained several outliers in each condition. Therefore, we used an M-estimator of central tendency, rather than the mean, to represent the "typical" threshold in each condition (Wilcox, 2005). Figure 4.6 shows the M-estimator of threshold (Huber's Ψ; Huber, 1981) measured in each group at each noise level on both days of testing.

A percentile bootstrap procedure (Wilcox, 2005, page 310) was used to assess group differences in the M-estimator of threshold at each level of noise on each day of testing. The familywise probability of a Type I error was set to $\alpha = 0.05$ for each session. None



(a)

(b)

Figure 4.6: Log-transformed threshold, expressed in terms of rms contrast, for each group in each condition on Day 1 (a) and Day 2 (b) of testing. Each bar represents the M-estimator of threshold measured with low (unfilled bars), medium (grey bars), or high (black bars) levels of external noise. M-estimators are robust measures of central tendency: in cases where the data are skewed, as was the case in the current experiment, an M-estimator is better than the mean at providing an estimate of a "typical" value (Wilcox, 2005). The particular M-estimator used here is Huber's Ψ. The error bars represent ±1 standard error of the M-estimator.

of the bootstrap tests was significant on Session 1 or Session 2. To compare thresholds in subjects who did and did not sleep between sessions, we combined thresholds from each of the three no-sleep groups into a single sleep group, and thresholds from the two no-sleep groups into a single no-sleep group. M-estimators of thresholds in the combined groups were then compared at each level of noise. Again, none of the comparisons was significant on either day. In summary, we did not obtain clear evidence that thresholds varied across the experimental groups, or between subjects that did or did not sleep between sessions, in Session 1 or 2.

The log-difference between thresholds in Sessions 1 and 2 was calculated for each subject in each condition. The M-estimator of the log-difference between thresholds was -0.27, -0.62, and -0.56 in the low, medium, and high external noise conditions, respectively. All three of these M-estimators differed significantly from zero ($p < .05$), which shows that thresholds decreased significantly at all levels of noise. To determine if the decrease in threshold varied across groups, a percentile bootstrap was used to compare the M-estimators measured for each group at each level of noise. None of the bootstrap tests was significant ($p > .05$). Next, the log-difference scores for subjects in the three no-sleep groups were combined into a single sleep group, and the scores from the remaining groups were combined into a single no-sleep group. Again, a percentile bootstrap performed on the M-estimators of the log-difference between thresholds found no significant ($p > .05$) difference between the combined sleep and no-sleep groups at any level of external noise. These analyses indicate that the difference between thresholds measured in the two sessions did not vary significantly across groups or between subjects that did or did not sleep.

In summary, we did not obtain clear evidence that thresholds in Session 1 and 2, or the difference between thresholds in the two sessions, varied across the experimental groups or was associated with the presence or absence of sleep. In this regard, these analyses of thresholds are consistent with the previous analyses on overall proportion correct.

4.5 Discussion

Using a 1-of-10 face identification task, we found that the proportion of correct responses increased by 0.17 and 0.1 during the first and second testing sessions, respectively. The difference between overall response accuracy in the two sessions was approximately 0.15, and therefore the magnitude of within-session and between-session learning effects was nearly the same. Between-session learning was significantly higher in groups that slept in-between sessions, but the effect of sleep on proportion correct was small (i.e., ≈ 0.026) and was restricted to performance in the first 105 trials in Session 2.

The finding that sleep had very small effects on learning differs from results obtained in several studies that used the texture discrimination task (e.g., Karni and Sagi 1991; Karni et al. 1994; Karni and Sagi 1993). Karni and Sagi (1991), for example, found that thresholds declined by approximately 22% between the first and second days of testing (see Figure 2 in Karni and Sagi, 1991). Censor et al. (2006) found that the size of sleep-related improvement in the texture discrimination task was a non-monotonic function of the number of trials used in the first session, with 26 trials per block (3-4 blocks per testing session) producing more overnight learning than 50 and 12 trials per block. Based on Censor et al.'s findings, it is tempting to attribute the small sleep effect found in the current experiment to the relatively large number of trials (i.e., 420) used in Session 1. However, using the same task and methods as in the current experiment, Hussain et al. (2003) found that reducing the number of practice trials in Session 1 does not result in greater amounts of between-session learning: during Session 2, subjects who received 420-840 trials during Session 1 performed significantly better than subjects who received received 21-210 trials during Session 1. Therefore we think it is unlikely that reducing the number of trials in Session 1 would increase the effect of sleep.

Mednick et al. (2002) reported that repeated, within-day testing on the texture discrimination task resulted in progressively higher thresholds (also see Mednick et al. 2003, 2005; but see Figure 2 in Stickgold et al., 2000a for a different result obtained with similar procedures). Between Sessions 1 and 2, for example, thresholds increased by 17% in Mednick et al. (2002) and 15% in Mednick et al. (2005). Some studies have also reported that performance deteriorates within a single session (Mednick et al., 2005; Ofen et al., 2007). Ofen et al., for example, found that performance with a single, above-threshold stimulus in the texture discrimination task decreased from 90% correct to 70% during the course of 8 blocks in a single test session (see Figure 4, in Ofen et al., 2007). In the current experiment, both No Sleep groups did exhibit a drop in performance at the start of Session 2, but the effect was small and was restricted to the first bin of 105 trials. Instead of decreasing, overall accuracy increased across sessions in both No Sleep groups (see Figure 4.2d). Furthermore, we found no evidence of performance decreasing within a session; in fact, response accuracy increased within each session in all groups.

The current experiment differs in several ways from previous studies that found larger effects of sleep or within-day deterioration. One potentially important difference concerns the psychophysical methods used to assess performance. Previous studies that measured learning with the texture discrimination task have obtained thresholds by adjusting the stimulus level using a variation of the traditional descending method of limits. With this method, increasing the number of trials per block, as was done by Censor et al. (2006) and Ofen et al. (2007), would necessarily reduce the variety of signal-to-noise ratios presented to the subject over the course of several trials or minutes. In contrast, the current experiment presented stimuli that varied significantly in terms of signal-to-noise ratio in a random order. This randomization procedure may minimize the adaptation that is thought to be important for generating within-session deterioration (Ofen et al., 2007).

Another obvious difference between studies is that the current experiment used a face identification task rather than the texture discrimination task. Face identification may depend on higher-level mechanisms that differ significantly from the mechanisms tapped by the texture discrimination task used in previous studies. It is possible, therefore, that learning in a face identification task differs qualitatively from the learning that is found with simpler perceptual tasks. However, results from other experiments are inconsistent with this hypothesis. For example, learning in a face identification task is specific for both the orientation and identity of the trained items (i.e., there is little generalization to new stimuli, Hussain et al., 2005). Furthermore, the benefits of face identification learning are long lasting, persisting for at least 9-18 months after the training sessions have ended (Hussain et al., 2007). Finally, face identification learning exhibits a fast within-session component and a slow between-session component (Figure 4.3), as has been found with simpler tasks (e.g., Karni and Sagi, 1993). Hence, the available evidence indicates that learning of the face identification task exhibits many of the characteristics of learning found with simpler perceptual tasks.

Nevertheless, the relative complexity of the neural network underlying face processing may alter the effects of sleep on learning. Sleep-dependent consolidation is thought to require the reactivation, during sleep, of cells engaged during the task. For example, consolidation of spatial learning has been linked to the reactivation of cells in the hippocampus (Wilson and McNaughton, 1994), and consolidation of visual conditioning is contingent on reactivation of visual cortex (Amzica et al., 1997). Indeed, some suggest that cortical activity in V1 is required during sleep for learning-related plasticity to be enabled (Jha et al., 2005). For the texture task, an imaging study indicates that V1 is the locus of practice effects (Schwartz et al., 2002), and the consolidation of such learning may depend on the reactivation of V1 during sleep. The learning we find with more complex stimuli may involve more than one cortical locus (such as IT), possibly diluting the effects of sleep if the entire network engaged during training is not reactivated during sleep.

Our results are consistent with the sustained improvements found in the absence of sleep for auditory and motor tasks, with mere passage of time after training (Gottselig et al., 2004; Robertson et al., 2004; Roth et al., 2005). In Roth et al. (2005), improvements on a verbal identification task emerged after at least six hours had elapsed. In Robertson et al. (2004), there was improvement on a finger-tapping task twelve hours after practice with no sleep between sessions. In Gottselig et al. (2004), restful waking, but not busy waking was equivalent to sleep in promoting learning of an auditory pattern discrimination task, suggesting that the key to consolidation might be the absence of interference from other tasks during the interim period (i.e., the interference hypothesis). However, performance on this auditory task also improved with no break between sessions, suggesting that the benefits from consolidation may be superfluous to those gained from continuous task performance. Likewise, face-identification requires little latent processing, as is clear from the performance of the 3hr-No sleep group, which improved despite the small time-window between sessions. These results are also consistent with a recent study showing negligible effects of sleep in face memory task (Sheth et al., 2008).

We found that within- and between-session learning were correlated positively for all groups (Figure 4.5). This finding is at odds with the results of Walker et al. (2003), who reported that sleep-dependent and within-session improvements on a motor task were uncorrelated. It should be pointed out, however, that although we found that the correlation was significant, within-session learning accounted for only a small proportion of the variance in between-session learning (i.e. $r^2 = 0.34^2 = 0.11$). It is unlikely that this relatively weak association is due to low reliability of our dependent measure, because the correlation between performance in Session 1 and 2 was 0.89 (Figure 4.2c). Instead, it appears that between-session learning depends substantially on factors that are not correlated with within-session learning. Stated this way, the current findings are similar to those reported by Walker et al. but, in addition, suggest that their conclusion that within- and between-session learning are mediated by "distinct and independent processes" (page 281, Walker et al., 2003) is too strong.

Overall, our results show that robust perceptual learning for a face identification task can be obtained in the absence of sleep, and that sleep has very little effect on between-session learning. The time-course of learning within each session was also unaffected by the presence or absence of sleep between sessions. These results are inconsistent with the notion that the bulk of perceptual learning is a latent, across-session phenomenon. It is commonly assumed that the latent gains are the basis of stimulus-specificity typically found in perceptual learning studies (Karni and Bertini, 1997; Karni et al., 1998; Karni and Sagi, 1993). The present findings, combined with other results from our lab indicating exemplar- and orientation- specificity of face learning (Hussain et al., 2005), suggest that stimulus-specificity of learning could just as well emerge from the improvements that occur within the training session.

References

Amzica, F., Neckelmann, D., Steriade, M., 1997. Instrumental conditioning of fast (20- to 50-hz) oscillations in corticothalamic networks. Proc Natl Acad Sci U S A 94 (5), 1985–1989.

Atienza, M., Cantero, J. L., Stickgold, R., 2004. Posttraining sleep enhances automaticity in perceptual discrimination. J Cogn Neurosci 16 (1), 53–64.

Brainard, D. H., 1997. The psychophysics toolbox. Spat Vis 10 (4), 433–436.

Brashers-Krug, T., Shadmehr, R., Bizzi, E., 1996. Consolidation in human motor memory. Nature 382 (6588), 252–255.

Censor, N., Karni, A., Sagi, D., 2006. A link between perceptual learning, adaptation and sleep. Vision Res 46 (23), 4071–4074.

Cohen, J., 1988. Statistical power analysis for the behavioral sciences, 2nd Edition. L. Erlbaum Associates, Hillsdale, N.J.

Fahle, M., 1994. Human pattern recognition: parallel processing and perceptual learning. Perception 23 (4), 411–427.

Fahle, M., Edelman, S., Poggio, T., 1995. Fast perceptual learning in hyperacuity. Vision Res 35 (21), 3003–3013.

Fenn, K. M., Nusbaum, H. C., Margoliash, D., 2003. Consolidation during sleep of perceptual learning of spoken language. Nature 425 (6958), 614–616.

Fiorentini, A., Berardi, N., 1981. Learning in grating waveform discrimination: specificity for orientation and spatial frequency. Vision Research 21 (7), 1149–1158.

Fischer, S., Hallschmid, M., Elsner, A. L., Born, J., 2002. Sleep forms memory for finger skills. Proc Natl Acad Sci U S A 99 (18), 11987–11991.

Gold, J., Bennett, P. J., Sekuler, A. B., 1999a. Identification of band-pass filtered letters and faces by human and ideal observers. Vision Res 39 (21), 3537–3560.

Gold, J., Bennett, P. J., Sekuler, A. B., 1999b. Signal but not noise changes with perceptual learning. Nature 402 (6758), 176–178.

Gottselig, J. M., Hofer-Tinguely, G., Borbely, A. A., Regel, S. J., Landolt, H.-P., Retey, J. V., Achermann, P., 2004. Sleep and rest facilitate auditory learning. Neuroscience 127 (3), 557–561.

Hauptmann, B., Reinhart, E., Brandt, S. A., Karni, A., 2005. The predictive value of the leveling off of within session performance for procedural memory consolidation. Brain Res Cogn Brain Res 24 (2), 181–189.

Horne, J. A., Ostberg, O., 1976. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. Int J Chronobiol 4 (2), 97–110.

Huber, P. J., 1981. Robust statistics. Wiley, New York.

Hussain, Z., Bennett, P., Sekuler, A., 2005. Perceptual learning of faces and textures is tuned to trained identities. Perception: ECVP Abstract Supplement 34.

Hussain, Z., Bennett, P., Sekuler, A., 2007. Superior identification of familiar visual patterns a year after learning. J Vision 7 (9), 797a.

Hussain, Z., Bennett, P. J., Sekuler, A. B., 2003. How much practice is needed to produce perceptual learning? J Vision 3 (9), 256a.

Jha, S. K., Jones, B. E., Coleman, T., Steinmetz, N., Law, C.-T., Griffin, G., Hawk, J., Dabbish, N., Kalatsky, V. A., Frank, M. G., 2005. Sleep-dependent plasticity requires cortical activity. J Neurosci 25 (40), 9266–9274.

Karni, A., Bertini, G., 1997. Learning perceptual skills: behavioral probes into adult cortical plasticity. Curr Opin Neurobiol 7 (4), 530–535.

Karni, A., Meyer, G., Rey-Hipolito, C., Jezzard, P., Adams, M. M., Turner, R., Ungerleider, L. G., 1998. The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. Proc Natl Acad Sci U S A 95 (3), 861–868.

Karni, A., Sagi, D., 1991. Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. Proc Natl Acad Sci U S A 88 (11), 4966–4970.

Karni, A., Sagi, D., 1993. The time course of learning a visual skill. Nature 365 (6443), 250–252.

Karni, A., Tanne, D., Rubenstein, B. S., Askenasy, J. J., Sagi, D., 1994. Dependence on rem sleep of overnight improvement of a perceptual skill. Science 265 (5172), 679–682.

McKee, S. P., Westheimer, G., 1978. Improvement in vernier acuity with practice. Perception and Psychophysics 24 (3), 258–262.

Mednick, S., Nakayama, K., Stickgold, R., 2003. Sleep-dependent learning: a nap is as good as a night. Nat Neurosci 6 (7), 697–698.

Mednick, S. C., Arman, A. C., Boynton, G. M., 2005. The time course and specificity of perceptual deterioration. Proc Natl Acad Sci U S A 102 (10), 3881–3885.

Mednick, S. C., Nakayama, K., Cantero, J. L., Atienza, M., Levin, A. A., Pathak, N., Stickgold, R., 2002. The restorative effect of naps on perceptual deterioration. Nat Neurosci 5 (7), 677–681.

Ofen, N., Moran, A., Sagi, D., 2007. Effects of trial repetition in texture discrimination. Vision Res 47 (8), 1094–1102.

Pelli, D. G., 1997. The videotoolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10 (4), 437–442.

Poggio, T., Fahle, M., Edelman, S., 1992. Fast perceptual learning in visual hyperacuity. Science 256 (5059), 1018–21.

Press, D. Z., Casement, M. D., Pascual-Leone, A., Robertson, E. M., 2005. The time course of off-line motor sequence learning. Brain Res Cogn Brain Res 25 (1), 375–378.

R Development Core Team, 2007. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (http://www.R-project.org).

Robertson, E. M., Pascual-Leone, A., Press, D. Z., 2004. Awareness modifies the skill-learning benefits of sleep. Curr Biol 14 (3), 208–212.

Roth, D. A.-E., Kishon-Rabin, L., Hildesheimer, M., Karni, A., 2005. A latent consolidation phase in auditory identification learning: time in the awake state is sufficient. Learn Mem 12 (2), 159–164.

Schoups, A. A., Vogels, R., Orban, G. A., 1995. Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. J Physiol 483 (Pt 3), 797–810.

Schwartz, S., Maquet, P., Frith, C., 2002. Neural correlates of perceptual learning: a functional mri study of visual texture discrimination. Proc Natl Acad Sci U S A 99 (26), 17137–17142.

Sheth, B., Nguyen, N., Janvelyan, D., 2008. How sleep influences our memory for faces. J Vision 8 (6), 876a.

Siegel, J. M., 2001. The rem sleep-memory consolidation hypothesis. Science 294 (5544), 1058–1063.

Song, S., Howard, J. H. J., Howard, D. V., 2007. Sleep does not benefit probabilistic motor sequence learning. J Neurosci 27 (46), 12475–12483.

Stickgold, R., James, L., Hobson, J. A., 2000a. Visual discrimination learning requires sleep after training. Nat Neurosci 3 (12), 1237–1238.

Stickgold, R., Whidbee, D., Schirmer, B., Patel, V., Hobson, J. A., 2000b. Visual discrimination task improvement: A multi-step process occurring during sleep. J Cogn Neurosci 12 (2), 246–254.

Tyler, C. W., Liu, H. C. L., McBride, B., Kontsevich, L., 1992. Bit-stealing: How to get 1786 or more gray levels for an 8-bit color monitor. In: Rogowitz, B. E. (Ed.), Proceedings of SPIE: Human vision, visual processing and digital display III. Vol. 1666. pp. 351–364.

Vertes, R. P., 2004. Memory consolidation in sleep; dream or reality. Neuron 44 (1), 135–148.

Vertes, R. P., Eastman, K. E., 2000. The case against memory consolidation in rem sleep. Behav Brain Sci 23 (6), 867–876.

Walker, M. P., Brakefield, T., Seidman, J., Morgan, A., Hobson, J. A., Stickgold, R., 2003. Sleep and the time course of motor skill learning. Learn Mem 10 (4), 275–284.

Walker, M. P., Stickgold, R., 2004. Sleep-dependent learning and memory consolidation. Neuron 44 (1), 121–133.

Wilcox, R. R., 2005. Introduction to robust estimation and hypothesis testing, 2nd Edition. Elsevier/Academic Press, Amsterdam.

Wilson, M. A., McNaughton, B. L., 1994. Reactivation of hippocampal ensemble memories during sleep. Science 265 (5172), 676–679.

Chapter 5

How much practice is needed to produce perceptual learning?

5.1 Abstract

We examined the amount of practice needed to improve performance on 10AFC face- and texture identification. On Day 1, subjects were grouped by amount of practice: a control group had zero trials of practice, and several experimental groups had practice that ranged from one to 40 trials per condition. On Day 2, all groups performed 40 trials per condition of the trained task. The effect of practice was estimated by comparing performance across groups on Day 2. In both tasks, increasing practice was associated with greater learning, but surprisingly small amounts of practice were required to improve performance. In the face identification task, for example, only one trial per condition on Day 1 was required to increase performance relative to the control group at the start of testing on Day 2. In the texture identification task, five trials per condition on Day 1 were required to increase performance relative to the control group. In both tasks, the advantage associated with small amounts of practice declined during the Day 2 session due to larger within-session learning in the control group.

Citation: Hussain, Z., Sekuler, A. B., & Bennett, P. J. How much practice is needed to produce perceptual learning? *Vision Research*, In Revision.

5.2 Introduction

Perceptual learning refers to improvements on sensory tasks brought about through practice (Ball and Sekuler, 1987; Fiorentini and Berardi, 1981; Matthews et al., 1999; Schoups et al., 1995; Sigman and Gilbert, 2000). The improvements - often stimulus-specific and long-lasting - are evidence for plasticity of the brain regions engaged by the sensory tasks (Karni and Bertini, 1997). There have been a number of investigations into the specificity (Ball and Sekuler, 1982; Sireteanu and Rettenbach, 1995; Furmanski and Engel, 2000; Sigman and Gilbert, 2000; Sowden et al., 2002; Yi et al., 2006; Hussain et al., In Press), and neural correlates of perceptual learning (Schoups et al., 2001; Schwartz et al., 2002; Rainer et al., 2004; Maertens and Pollmann, 2005; Kourtzi et al., 2005; Raiguel et al., 2006; Pourtois et al., 2008; Yotsumoto et al., 2008), but the amount of practice needed to produce learning has been rarely been examined directly (Wright and Sabin, 2007).

How much practice is needed to produce the long-lasting and stimulus-specific effects that are the hallmark of perceptual learning? In the auditory domain, one study directly examined how much practice is needed to produce learning in temporal and frequency discrimination tasks (Wright and Sabin, 2007). Participants were given either 360 or 900 practice trials per day for six days. Temporal discrimination improved with 360 trials per day, but frequency discrimination did not, a result that suggests that a critical amount of practice, which varies across tasks, may be needed for learning to occur. Consistent with this idea, Hauptmann et al. (2005) showed that between-session improvements on a letter enumeration task do not emerge unless subjects have reached asymptote within the first training session. Finally, the critical amount of practice also may depend subtle aspects of the experimental procedure: For example, Hawkey et al. (2004) – using a task that differed from the one used by Wright and Sabin – found that stimulus-specific improvements in frequency discrimination could be obtained with fewer than 200 practice trials.

Thus, although the few studies addressing the issue suggest that there may be critical levels of practice, the amount needed to improve performance on the remaining variety of perceptual tasks, particularly in the visual domain, is not known. In the visual domain, improvements in perception are frequently measured after providing observers with extensive practice, up to and even exceeding 4000-10,000 trials over the course of 2 to 18 days (Gold et al., 1999b; Furmanski and Engel, 2000; Sigman and Gilbert, 2000; Gold et al., 2004; Chung et al., 2006; Richards et al., 2006; Husk et al., 2007). Large amounts of practice reveal within-session improvements and ensure asymptotic performance, but do not address whether performance improves when practice is restricted to a few trials. Here, we ask whether small amounts of practice can elicit improvements on two tasks known to be amenable to the effects of training, and for which stimulus-specific effects of practice previously have been shown: 10-AFC texture identification and face identification (Gold et al., 1999b, 2004; Hussain et al., In Press, 2005). These tasks enable the comparison of learning of frequently encountered objects versus novel objects, where both object classes comprise multiple features that differentiate exemplars within the class.

Table 5.1: Number of subjects run in each practice condition in the face- and texture identification tasks.

Trials per condition	Face	Texture
0	20	23
1	25	21
5	19	22
10	20	16
20	10	-
40	27	28

5.3 Methods

5.3.1 Subjects

Two hundred and thirty-one McMaster University undergraduate students participated in this experiment (mean age = 21.21, SD = 3.66). All subjects had normal or corrected-to-normal Snellen visual acuity. All subjects received a small stipend ($10/hour) or partial course credit for participating in the experiment, and were naive with respect to the task.

5.3.2 Apparatus and Stimuli

Stimuli were generated in Matlab (v. 5.2) using the Psychophysics and Video Toolboxes (Brainard, 1997; Pelli, 1997). Stimuli were displayed on a 21" Sony Trinitron monitor with 1024 x 768 pixels at a frame rate of 85 Hz. Average luminance was 73.12 cd/m^2. The monitor calibration data were used to build a 1779-element lookup table (Tyler et al., 1992) and customized computer software constructed the stimuli on each trial by selecting the appropriate luminance values from the calibrated lookup table and storing them in the display's eight-bit lookup table.

The face stimuli were faces of five males and five female faces cropped to show only internal features. All of the faces had the same global amplitude spectrum, see Gold et al. (1999a) for a more detailed account of the stimuli. The textures were band-limited noise patterns created by applying an isotropic, band-pass (2-4 cy/image) ideal spatial frequency filter to gaussian noise (see Figure 1). The faces and textures subtended 4.8 x 4.8 deg of visual angle from the viewing distance of 114 cm. During the experiment, stimulus contrast was varied across trials using the method of constant stimuli. Seven levels of contrast were spaced equally on a logarithmic scale, and spanned a range that was sufficient to produce significant changes in performance in virtually all subjects. The stimuli were shown in one of three levels (low, medium and high) of static two-dimensional Gaussian noise, created by sampling from distributions with variances of .001, .01, and .1. Hence, subjects viewed each stimulus at a signal-to-noise ratio that varied significantly across trials. Thus there were 21 different stimulus conditions (seven contrast levels x three external noise levels). The contrasts used for the textures at the low noise level were 4 times higher than the contrasts used for the faces.

5.3.3 Procedure

Each subject was tested on two consecutive days, at roughly the same time each day. On Day 1, observers were assigned to one of the practice groups (see Table 5.1). The different practice groups saw either 0, 1, 5, 10, 20 or 40 trials per stimulus condition (henceforth referred to as the 1-, 5-, 10-, 20- and 40-trials groups) which corresponds to a total of 0, 21, 105, 210, 420 and 840 trials per session. The 20-trials condition was only run for the face identification task. Importantly, the average stimulus contrast presented at each noise level was constant across the different practice conditions. On Day 2, all subjects were tested in the 40 trials/condition (i.e., a total of 840 trials).

All subjects were seated in a darkened room 114 cm away from the monitor. Viewing was binocular, and viewing position and distance were stabilized with an adjustable chin-rest. The experiment started after a 60-second period during which the subject adapted to the average luminance of the display. A trial began with the presentation of a fixation point in the center of the screen for 100 ms (black high-contrast spot; 0.15 x 0.15 deg), followed by a randomly selected face/texture, presented for approximately 200 ms at the center of the screen in one of 21 stimulus conditions. After the face/texture disappeared,

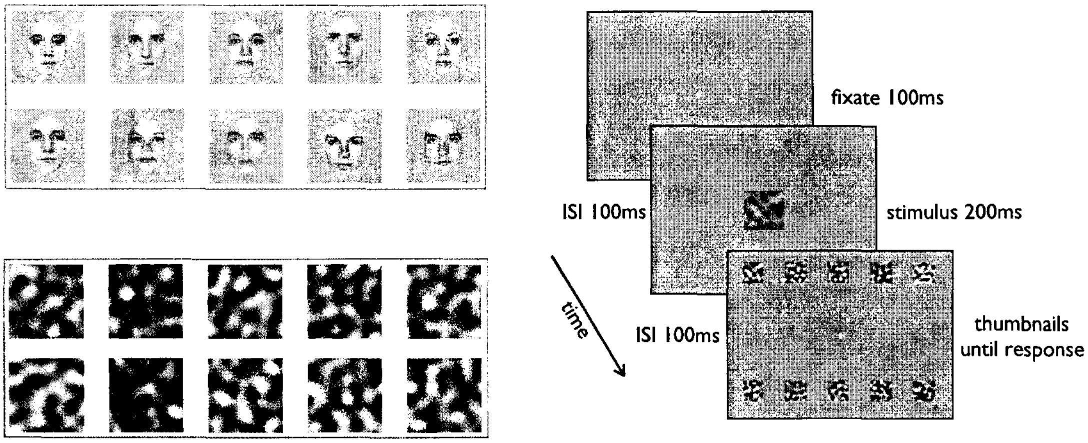


Figure 5.1: Stimuli used in the face and texture identification tasks, and a schematic representation of the trial sequence in the identification task.

the entire set of 10 faces/textures was presented as noiseless, high-contrast thumbnail images, each subtending 1.7 x 1.7 deg. Five thumbnails were presented on the top half of the screen, and five on the bottom half. The subject's task was to decide which one of the 10 faces/textures had been presented during the trial, and to respond by clicking on the chosen face/texture. The location of each face/texture in the response window was constant across subjects, trials, and sessions. Auditory feedback in the form of high-pitched (correct) and low-pitched (incorrect) tones informed the subject about the accuracy of each response, and the next trial began one second after presentation of the feedback. Figure 1 shows a schematic illustration of the task.

The duration of the practice session (Day 1) for the 40-trials group was approximately one hour, and correspondingly shorter in the other groups. To equate time spent in the laboratory, subjects who received fewer than 40 trials per condition performed an additional task after the completion of the experimental task. The additional task measured the accuracy of memory for the orientation of a high-contrast line and was designed to differ significantly from the face- and texture-identification tasks (see Bennett et al., 2007, page 803, for a detailed description of this task). Subjects performed the orientation memory task until the total duration of the experimental session was approximately one hour, and therefore the total time spent in the laboratory on Day 1 was equated across groups.

5.4 Results

Statistical analyses were performed in R (R Development Core Team, 2007). Multiple comparisons were done using the R package `multcomp` (Hothorn et al., 2008). Effect size is expressed in terms of Cohen's f (Cohen, 1988).

5.4.1 40-trials Groups

We first examined whether 40 trials per condition were sufficient to produce significant learning. The responses obtained from the 40-trials groups on each day were divided into eight bins of 105 trials each (Figure 5.2). Proportion correct in the texture and face groups was then analyzed separately with a 2 (Day) x 8 (Bin) within-subjects analysis of variance (ANOVA). In the texture identification task, the ANOVA revealed significant main effects of Day, $F(1,27) = 123$, $p < .0001$, $f = 0.52$, and Bin, $F(7,189) = 50.56$, $p < .0001$, $f = 0.88$, and a significant Day × Bin interaction, $F(7,189) = 7.54$, $< .0001$, $f = 0.32$. The interaction was analyzed by evaluating the simple main effect of Day at each Bin. The difference between accuracy on Days 2 and 1 was largest in Bin 1 ($D = 0.29$, $CI_{95\%} = [0.23,\ 0.35]$) and declined to an average of 0.17, $CI_{95\%} = [0.14,\ 0.19]$, in Bins 6-8. Nevertheless, response accuracy measured on Day 2 was higher than on Day 1 in all Bins ($t(27) \geq 6.45$, $p < .0001$ in all cases). These results also suggest that there was more within-session learning for the 40-trials group on Day 1 than on Day 2.

In the face identification task, the ANOVA found significant main effects of Day, $F(1,26) = 254$, $p < .0001$, $f = .76$, and Bin, $F(7,182) = 28.5$, $p < .0001$, $f = .67$, as well as a significant Day × Bin intertaction, $F(7,182) = 6.97$, $p < .0001$, $f = .31$. The difference between days was largest in the initial bin ($D = 0.26$, $CI_{95\%} = [0.21,\ 0.31]$) and declined to an average of 0.13, $CI_{95\%} = [0.11,\ 0.15]$ in the last three bins. Nevertheless, as was the case with textures, the simple main effect of Day was significant at each Bin ($t(26) \geq 5.42$, $p < .0001$ in all cases). Also, the analyses suggest that there was more within-session learning on Day 1 than on Day 2 for this group.

Inspection of Figure 5.2 shows that average response accuracy on Day 1 was significantly greater in the face condition than in the texture condition ($CI_{95\%} = [0.02,\ 0.16]$, $t(53) = 2.49$, $p = 0.015$). Average performance with faces also was better than with textures on Day 2, but the difference between groups was not statistically significant ($CI_{95\%} = [-.03,\ .14]$, $t(53) = 1.24$, $p = 0.22$).

The current results are consistent with previous reports that 40 trials per condition on Day 1 are sufficient to produce learning in these texture and face identification tasks (Hussain et al., 2005, 2009, In Press).

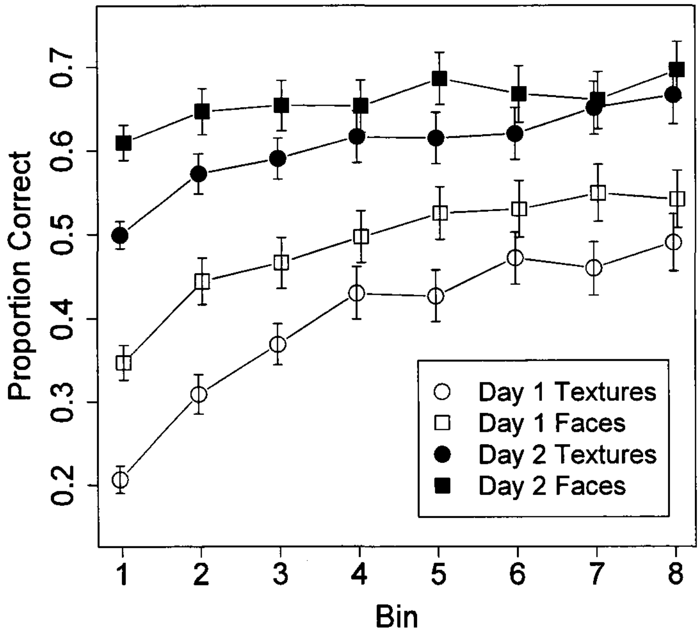


Figure 5.2: Proportion correct on Days 1 and 2 for the 40-trials groups in the face- and texture-identification tasks. Data from each day have been divided into eight bins of 105 trials each. The symbols represent average proportion correct. Error bars represent ±1 SEM.

5.4.2 Comparison of Groups on Day 2

In this section we compare response accuracy measured in all groups on Day 2. The analyses addressed whether exposure to textures or faces on Day 1 improved performance relative to the zero-trials groups, and whether groups that received 1-20 trials per condition performed worse than the 40-trials groups.

5.4.2.1 Texture identification

Responses obtained from subjects in the texture identification task were divided into eight bins of 105 trials each. Proportion correct in each bin is plotted as a function of bin number in Figure 5.3. Note that the bins are numbered 9-16 to differentiate them from the eight bins of trials presented to the 40-trials group on Day 1. Figure 5.3 clearly shows that the 10- and 40-trials groups performed better than the zero-trials group. The evidence of learning in the 5-trials group is mixed: response accuracy in that group was higher than in the zero-trials group in Bins 9-12, but there was no difference between groups in Bins 13-16. There was no evidence of learning in the 1-trial group, and therefore the zero-trials and 1-trial group were pooled into a single baseline group in some of the following analyses to increase statistical power.

Response accuracy was analyzed with a 5 (Group) × 8 (Bin) ANOVA. The main effects of Bin, $F(7,735) = 144$, $p < .0001$, $f = 1.02$, and Group, $F(4,105) = 4.84$, $p = .0012$, $f = 0.13$, were significant, as was the Group × Bin interaction, $F(28,735) = 3.11$, $p < .0001$, $f = 0.25$. The main effect of Bin reflects the fact that proportion correct increased across Bins in every group. The main effect of Group is illustrated in the bottom part of Figure 5.3, which shows proportion correct averaged across bins for each Group. Pairwise differences between groups were evaluated using the method described by Westfall (1997), which adjusts alpha to control for family-wise error: Only differences between the 40-trials group and the zero-, 1-, and 5-trials group were significant (adjusted-$p < .02$).

The Group × Bin interaction was analyzed by evaluating the simple main effect of Group at each Bin, while using the while using the Holm-Bonferroni adjustment (Holm, 1979) to maintain a familywise Type I error rate of .05 across all eight tests. The simple main effect of Group was significant at only Bins 9-12 ($F(4,105) \geq 3.91$, adjusted-$p \leq .026$, $f \geq 0.33$). Next, each significant simple main effect of Group was analyzed with five contrasts that tested the hypotheses that i) response accuracies in each of the 40-, 10-, and 5-trials groups was greater than accuracy in a baseline group that was created by pooling the zero- and 1-trials groups; and ii) accuracies in both the 5-, and 10-trials groups were less than accuracy in the 40-trials group. For each set of five contrasts, the familywise Type I error rate was set to .05 (Westfall, 1997). In Bin 9 (i.e., the first bin of trials on Day 2), all contrasts were significant: response accuracy in each of the 40-, 10-, and 5-trials groups was greater than in the baseline group, and accuracy in the 5-, and 10-trials groups was less than the 40-trials group. The results in Bin 10 were the same as in Bin 9, except that the 5-trials group did not differ significantly from the baseline group. The results in Bin 11 were the same as in Bin 10, except that the 10-trials group did not differ from the baseline group. Hence, in Bin 11 response accuracy in the 40-trials group was significantly greater than in the baseline, 5-trials, and 10-trials groups. Finally, in Bin 12 the 40-trials group was significantly greater than the baseline and 5-trials groups.

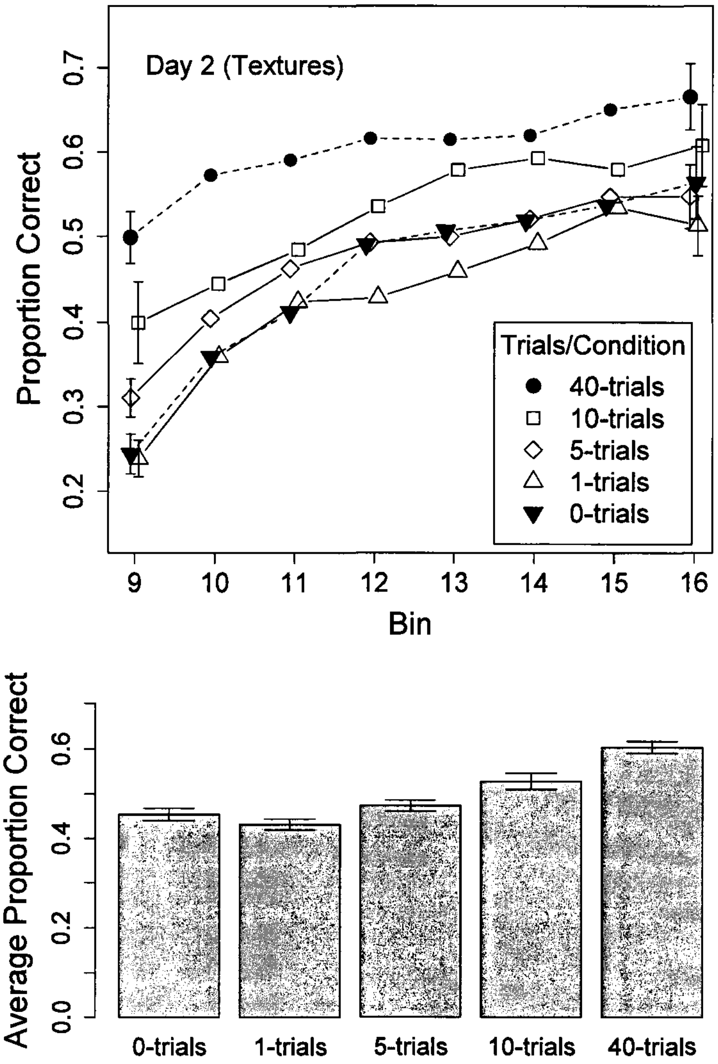


Figure 5.3: Top: Proportion correct on Day 2 for all groups tested with textures. The data have been divided into eight bins of 105 trials each. The bins are numbered 9-16 to differentiate them from the eight bins presented on Day 1. The symbols represent average proportion correct. Error bars represent ±1 SEM. Standard errors were nearly constant across bins. For clarity, therefore, only error bars in Bins 9 and 16 are shown. Bottom: Proportion correct – averaged first across all trials and then across subjects – on Day 2. Error bars represent ±1 SEM.

Within-session learning from Bins 9-12 was compared between the baseline group and the 5-, 10- and 40-trials practice groups combined, all of which differed from the baseline group at least in Bin 9. On average, the 5-, 10- and 40-trials groups improved by 14% from Bin 9 to Bin 12, whereas the baseline group improved by 22%. The 8% difference in amount of within-session improvement between the practiced and baseline groups was significant ($t(78.238) = 3.3089, p = .001$).

These analyses suggest that 40-, 10- and even 5 trials of practice per condition on Day 1 were sufficient to increase response accuracy relative to subjects who received zero or one trial of practice per condition, at least at the start of the testing session on Day 2. However, larger amounts of within-session learning in the baseline groups reduced that advantage by the end of Day 2.

5.4.2.2 Face identification

Proportion correct from subjects in the face identification task is plotted as a function of bin number in Figure 5.4. Figure 5.4 clearly shows that there was virtually no difference in performance between the 20- and 40-trials groups, and that both groups performed significantly better than the zero-trials group. There also is some suggestion that performance in the 1-, 5-, and 10-trials groups was worse than the 40-trials group but better than the zero-trials group, at least at the beginning of the test session. A 6 (Group) x 8 (Bin) ANOVA found significant main effects of Bin, $F(7, 805) = 58.85$, $p < .0001$, $f = 0.65$, and Group, $F(5, 115) = 5.01$, $p = .0003$. $f = 0.14$, and a significant Group × Bin interaction, $F(35, 805) = 2.85$, $p < .0001$, $f = 0.26$. As was the case with textures, the main effect of Bin reflects the fact that proportion correct increased across Bins in every group. The main effect of Group is illustrated in the bottom part of Figure 5.4, which shows proportion correct averaged across bins for each Group. Pairwise differences between groups were evaluated using the method described by Westfall (1997). Average response accuracy in the zero-trials group differed from accuracy in the 40- and 20-trials groups. In addition, the 1- and 10-trials groups both differed from the 40-trials group. Differences between the 20-trials group and the 1- and 10-trials groups approached significance ($p = .06$).

The Group × Bin interaction was analyzed by evaluating the simple main effect of Group at each Bin, using the Holm-Bonferroni adjustment (Holm, 1979) to control familywise Type I error rate. The simple main effect of Group was significant at all bins except 14 and 15 ($F(4, 105) \geq 2.98$, adjusted-$p \leq .043$, $f \geq 0.29$). Each significant simple main effect of Group was analyzed with nine contrasts that tested the hypotheses that i) response accuracies in each of the 1-40 trials groups was greater than accuracy in the zero-trials group; and ii) accuracies in each of the 1-20 trials groups was less than accuracy in the 40-trials group. Familywise Type I error rate for each set of contrasts was controlled using the method described by Westfall (1997). In Bin 9, all of the contrasts were significant (adjusted-$p \leq .017$) except for the one comparing accuracy in the 20- and 40-trials groups. The same results were obtained in Bin 10, except that the 1- and 5-trials groups no longer differed significantly from the zero-trials group (adjusted-$p = .06$ in both cases). The results in Bin 11-13 were the same as in Bin 10, except the difference between the 5- and 40-trials groups was not significant. Finally, in Bin 16 the only significant differences were between the 40-trials group and the 0- and 10-trials groups.

Within-session learning from Bins 9-13 was compared between the zero-trials group and the 1-, 5- and 10-trials groups combined. On average, the 1-, 5- and 10-trials groups improved by 13% from Bin 9 to Bin 13, whereas the zero-trials group improved by 21%. The 7% difference in amount of within-session improvement between the practiced and zero-trials groups was significant ($t(33.626) = 2.822, p = .007$).

These analyses suggest that all groups that received some practice (even the most minimal levels), with face stimuli on Day 1 performed better than the zero-trials group, at least at the start of the testing session on Day 2. However, larger amounts of within-session learning by the zero-trials group from bins 9-13 reduced the advantage of the practiced groups.

5.4.3 Contrast thresholds

Psychometric functions were fit to the data from individual subjects on Day 2 to calculate identification thresholds, defined as the rms contrast needed to attain 50% correct. The experiments used the same fixed set of stimulus contrasts for all subjects, and therefore reliable thresholds could not be obtained for every subject. Furthermore, the fact that we were studying learning meant that subjects did not receive practice to stabilize performance. Consequently, approximately 3% of the thresholds – or 14 of 462 in the texture identification conditions, and 19 of 534 in the face identification conditions – were impossible values (i.e., rms contrasts less than 0). Even after removing these values, the log-transformed thresholds contained outliers and were strongly positively skewed in each condition. For data exhibiting these characteristics, the Modified One-step M-estimator (MOM) is a better index of a typical score than the sample mean. Furthermore,

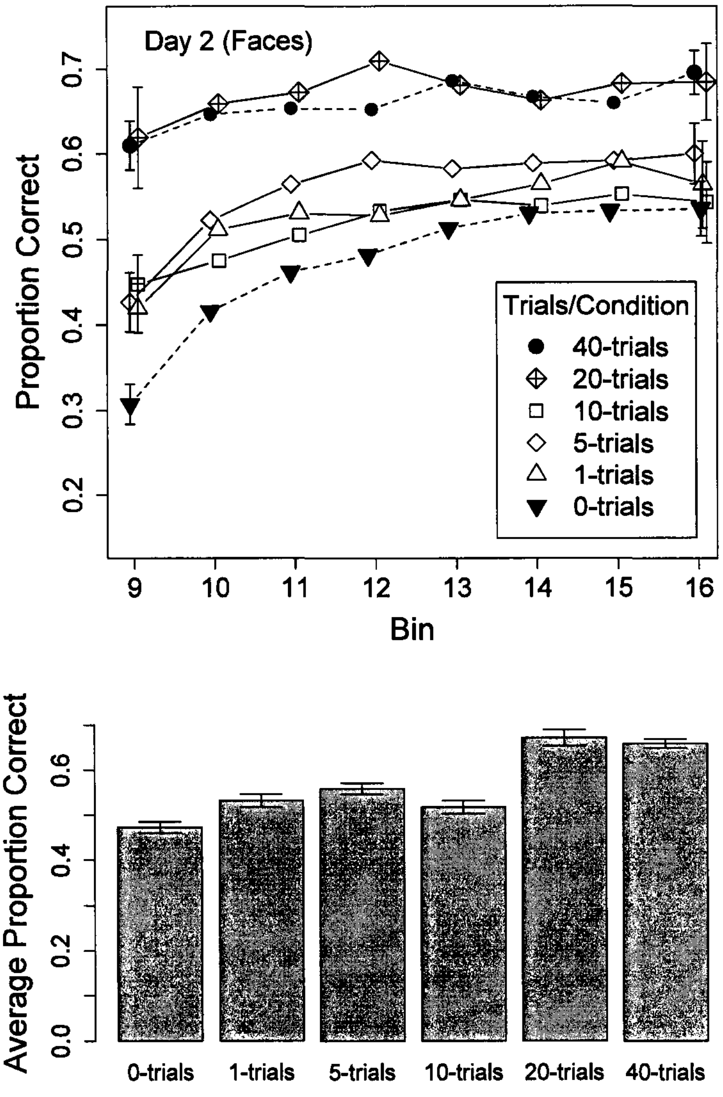


Figure 5.4: Top: Proportion correct on Day 2 for all groups tested with faces. Bottom: Proportion correct averaged across all trials on Day 2. Plotting conventions are the same as in Figure 5.3.

analyses of group MOMs are more sensitive than standard ANOVA methods when the data are skewed and contain outliers (Wilcox, 2005). Therefore, the following analyses were conducted on the MOMs of the log-transformed thresholds in each condition.

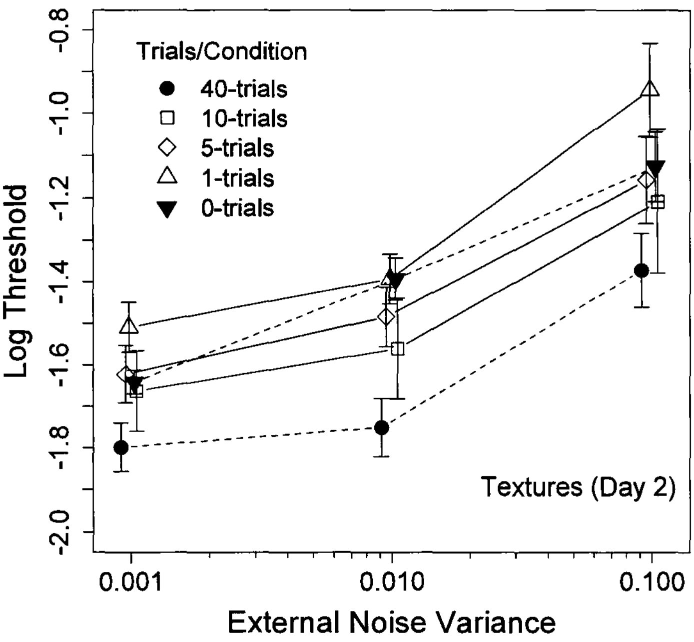


Figure 5.5: Texture identification thresholds measured on Day 2 plotted as a function of external noise. Threshold was defined as the rms contrast needed to attain 50% correct responses. The same levels of external noise were used in all conditions: the symbols are displaced, slightly, along the horizontal axis to make it easier to discriminate among conditions. Error bars represent ±1 standard error.

Texture identification thresholds on Day 2 are shown in Figure 5.5. The threshold-vs-noise curves are qualitatively similar to those obtained in previous studies (Gold et al., 1999b, 2004). A percentile-bootstrap method (Wilcox, 2005, page 368) found that the MOM of threshold – averaged across the three noise levels – varied significantly across groups ($p = .01$), and that MOMs, averaged across groups, varied significantly across noise levels ($p < .001$). Hence, the main effects of Group and Noise were significant. However, the Group × Noise interaction was not significant ($p = 0.52$). These results are consistent with the idea that learning shifted the threshold-vs-noise curve vertically in the log-log plot (Gold et al., 1999b, 2004). The main effect of Group was analyzed by conducting multiple, pairwise comparisons, which found that thresholds in the 40-trials group were significantly lower than thresholds in the 0-trials and 1-trial groups ($p < .05$).

Face identification thresholds on Day 2 are shown in Figure 5.6. As in the texture conditions, face identification thresholds increased with increasing levels of noise. However, the quadratic component of the threshold-vs-noise curve is less noticeable in the face conditions than in the texture conditions. As was the case with texture identification thresholds, a percentile-bootstrap method (Wilcox, 2005, page 368) found significant main effects of Group ($p = .02$) and Noise ($p < .001$), but the Group × Noise interaction was not significant ($p = .31$). Hence, the results are consistent with the hypothesis that learning shifted face identification thresholds vertically in the log-log plot (Gold et al., 1999b, 2004). The main effect of Group was analyzed by conducting multiple, pairwise comparisons, which found that thresholds in the 40- and 20-trials groups were both significantly lower than thresholds in the 0-trials group, and that thresholds in the 40-trials group also were lower than thresholds in the 1-trial group ($p < .05$).

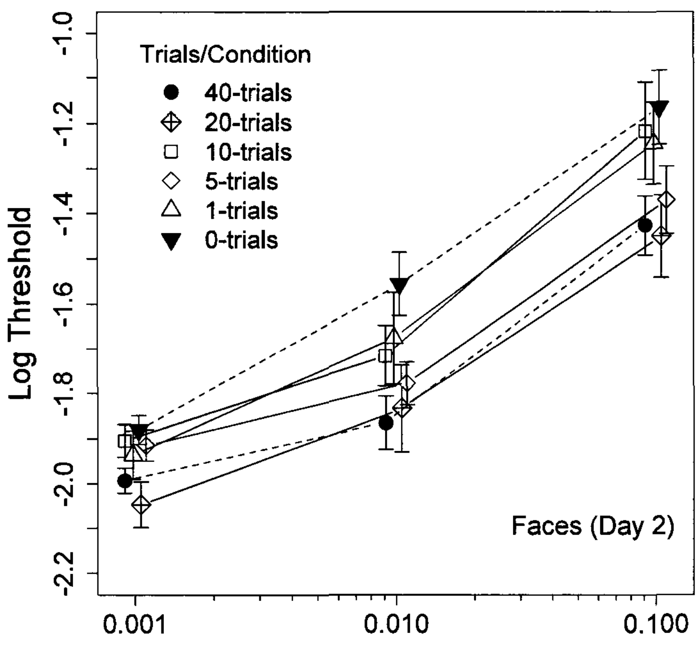


Figure 5.6: Face identification thresholds measured on Day 2 plotted as a function of external noise. Threshold was defined as the rms contrast needed to attain 50% correct responses. The same levels of external noise were used in all conditions: the symbols are displaced, slightly, along the horizontal axis to make it easier to discriminate among conditions. Error bars represent ±1 standard error.

5.5 Discussion

Small amounts of practice on texture- and face identification on Day 1 benefitted accuracy at the start of the session on Day 2 relative to a zero-trials group who did not practice the identification task. For face identification, there was an advantage with just one practice trial per condition (i.e., practice with only 21 trials on Day 1). The best performance on Day 2, across all bins, both with faces and textures, was achieved with larger amounts of practice (i.e., 40 trials per condition, a total of 840 trials on Day 1). The benefits of small amounts of practice were diminished in the latter part of the session due to substantial within-session learning by the zero-trials groups. The results point to the importance of within-session learning in the overall gains with practice on texture- and face identification; for these tasks, more practice clearly is better even though some improvements can be detected after limited amounts of practice.

The benefits of small amounts of practice on accuracy were not reflected in contrast thresholds. This failure to find a difference in thresholds is probably due to the fact that small amounts of practice benefited performance only at the start of the Day 2 session, whereas the thresholds were based on all trials in the session. An adaptive procedure (e.g., staircase), rather than the method of constant stimuli may have been better suited to capture the early effects of training on thresholds. However, even with the methods used in the current experiments, thresholds in the 40 - and 20-trials groups in the face identification task were equivalent on Day 2, indicating that substantial threshold reductions can be obtained with a fraction of the amount of practice used in other studies of face learning (Gold et al., 1999b, 2004).

Overall, the critical amount of practice needed to raise performance is larger for texture- than for face identification, a result that complements data from the auditory domain showing differences across tasks in the critical amount of practice needed to improve performance (Wright and Sabin, 2007). In the next section we discuss what might account for differences in the critical amount of practice needed for faces versus textures

5.5.1 Face- versus texture identification

One trial per condition raised performance on face- but not texture-identification. This may have been due to differences in task difficulty arising from the unfamiliar quality of the texture stimuli, which required observers to develop an identification strategy for the textures. Two aspects of the data support this possibility: i) accuracy in all bins on Day 1 was lower in the texture- than in the face identification task (Figure 2); ii) accuracy of all practice groups in the face identification task, including the 1-trial group, was equivalent in Bin 1 on Day 1, whereas accuracy of the 1-trial group in the texture identification task, was lower than that of the other groups, see Figure 5.7. Classification images have shown that subjects consistently use the same stimulus regions to distinguish faces, whereas the regions used with textures are idiosyncratic across observers (Gold et al., 2004), which confirms the idea that observers must develop identification strategies for textures but not for faces. We suggest that brief exposure facilitates learning when the stimulus properties are known in advance, as was the case with faces, but not with textures. Future studies investigating the role of differential strategy on the minimum number of trials should ensure that task difficulty is equated across tasks.

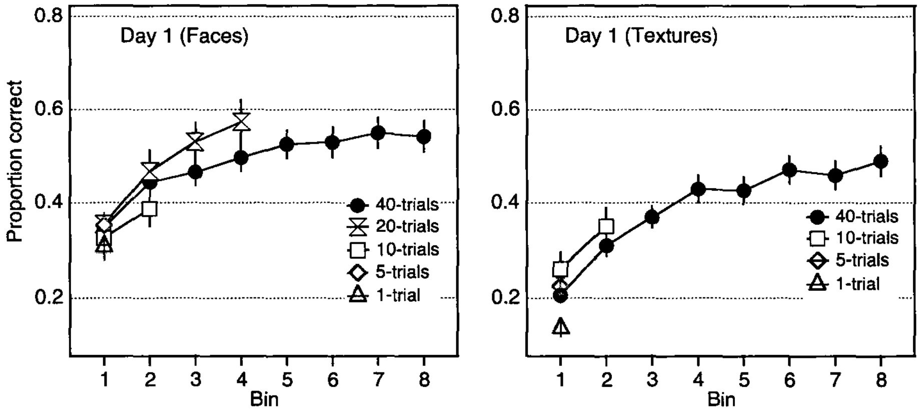


Figure 5.7: Proportion correct on Day 1 for all groups tested with faces and textures. There was no group difference in Bin 1 for faces ($F(4, 96) = .48, p = .75$), but there was a group difference in Bin 1 for textures ($F(3, 83) = 5.55, p = .001$). A TukeyHSD test confirmed that the 1-trial group differed from the other groups in Bin 1, and that none of the other groups differed from each other.

In other respects, we have elsewhere found perceptual learning of face- and texture-identification to be remarkably similar: in both cases learning is exemplar-specific (Hussain et al., 2005), orientation-specific (Hussain et al., 2009, In Press) and long-lasting (Hussain et al., 2007). As well, both tasks exhibit partial transfer of learning to novel stimuli when the time course of learning is examined at a fine scale. The only other difference between the two tasks is that there is a small amount of transfer of learning across orientation with faces but none with textures (Hussain et al., In Press), which we have speculated as arising from differences in the structural regularities between the stimuli (i.e., average spatial structure in objects like faces versus the lack thereof in objects like textures). Therefore the structural regularities in objects, aside from facilitating transfer of learning across conditions, might also hasten learning relative to objects without this property.

5.5.2 Within-session learning versus perceptual deterioration

Interestingly, some studies have suggested that it is possible to receive *too much* practice on certain tasks: increasing the number of practice trials per day beyond some upper limit either leads to no further improvement on certain tasks (Ofen-Noy et al., 2003; Savion-Lemieux and Penhune, 2005; Wright and Sabin, 2007) or it interferes with learning (Censor et al., 2006; Mednick et al., 2005; Ofen et al., 2007). For example, performance on a texture discrimination task has been reported to get worse when practice sessions comprising large numbers of trials are conducted within the same day, unless subjects rest or sleep between sessions (Mednick et al., 2005). The tendency for practice to worsen performance within the same session (aka perceptual deterioration) is thought to be due to an adaptation-like process dependent on the frequency of repetitions of the same type of trial (Censor et al., 2006; Ofen et al., 2007). Perceptual deterioration is in marked contrast to the within-session improvements reported for a number of tasks (Fiorentini and Berardi, 1981; Poggio et al., 1992; Karni and Sagi, 1993; Beard et al., 1995; Matthews et al., 1999), and to the robust-within session learning found here. Indeed, there was no evidence of perceptual deterioration in the current tasks, either with small or large amounts of practice within the session. Instead, performance steadily increased during the session. The factors that produce deterioration rather than learning within a session need to be clarified.

5.5.3 Amount versus distribution of practice

Some have suggested that distributed practice is more beneficial for perceptual and motor learning than massed practice (Savion-Lemieux and Penhune, 2005; Ofen-Noy et al., 2003), consistent with research on learning in memory tasks (Woodsworth and Schlosberg, 1956). Ofen-Noy et al. (2003) showed that increasing the amount of training from four to eight blocks did not enhance performance on a mirror reading task, and that the data were better described by a power function constrained by the time schedule of training, than by a function constrained by the number of trial repetitions. Savion-Lemieux and Penhune (2005) also showed that distribution of practice over several days was the critical factor in learning and retention of a timed motor sequence task. Here, many practice trials closely spaced within the training session resulted in better performance on Day 2 relative to groups that received fewer practice trials. Furthermore, an additional analysis showed no group difference in average accuracy of 840 trials conducted within a session versus 840 trials conducted across the two sessions (i.e., 40- and zero-trials groups versus the remaining groups), which confirmed that there was no advantage from a gap between sessions. We suggest that for certain tasks, overall performance is determined by the sheer amount practice than by the distribution of practice across time.

5.5.4 Rapid learning versus one-trial learning

Rapid improvements in performance have been shown in the time-course of learning for several visual tasks, usually within the first 100 trials on the first day of training on the task (Poggio et al., 1992; Fahle et al., 1995; Karni and Sagi, 1993; Tovee et al., 1996; Sireteanu and Rettenbach, 1995; Beard et al., 1995). These rapid improvements are thought to reflect procedural learning that differs from the stimulus-specific gains that arise from better representation of the relevant stimulus properties (Karni and Bertini, 1997). Yet, highly specific learning can be obtained with a single target exposure, (i.e., the phenomenon of one-trial learning; Sahley et al., 1981). One-trial learning typically has been demonstrated with animals in the context of associative learning (Chang and Gelperin, 1980; Balderrama, 1980; Sahley et al., 1981; Brandon and Coss, 1982; Malin et al., 1986; Armstrong et al., 2006; Cook and Fagot, 2009), but there are examples of one-trial learning in humans as well (Rozin, 1986; Taieb-Maimon, 2007). Indeed, some researchers have shown that stimulus-specific perceptual learning can occur early in training, within the first 200 trials (Fahle et al., 1995; Hawkey et al., 2004). In the current tasks, there was rapid within-session learning, as well as a benefit from one trial per condition for face identification. We have elsewhere shown that learning of texture- and face identification is stimulus-specific (Hussain et al., 2005, 2009, In Press); whether stimulus-specificity of learning of these tasks can emerge from brief practice is an open question.

5.5.5 Conclusions

Small amounts of practice can improve accuracy on texture- and face identification, but large amounts of practice are better. Faces but not textures can be learned with as little as one trial per condition. The benefits from small amounts of practice are evident early in the test session, but more difficult to discern in average performance pooled over many trials due to the substantial influence of within-session learning by the relatively naive groups. Performance does not deteriorate across trials for these tasks. Instead, learning is an incremental function of practice.

References

Armstrong, C., DeVito, L., Cleland, T., 2006. One-trial associative odor learning in neonatal mice. Chemical Senses 31 (4), 343–349.

Balderrama, N., 1980. One trial learning in the american cockroach, perplaneta-americana. Journal of Insect Physiology 26 (8), 499–504.

Ball, K., Sekuler, R., 1982. A specific and enduring improvement in visual-motion discrimination. Science 218 (4573), 697–698.

Ball, K., Sekuler, R., 1987. Direction-specific improvement in motion discrimination. Vision Res 27 (6), 953–965.

Beard, B., Levi, D., Reich, L., 1995. Perceptual learning in parafoveal vision. Vision Research 35 (12), 914–924.

Bennett, P. J., Sekuler, R., Sekuler, A. B., 2007. The effects of aging on motion detection and direction identification. Vision Res 47 (6), 799–809.

Brainard, D. H., 1997. The psychophysics toolbox. Spat Vis 10 (4), 433–436.

Brandon, J., Coss, R., 1982. Rapid dendritic spine stem shortening during one-trial learning - the honebees 1st orientation flight. Brain Research 252 (1), 51–61.

Censor, N., Karni, A., Sagi, D., 2006. A link between perceptual learning, adaptation and sleep. Vision Res 46 (23), 4071–4074.

Chang, J., Gelperin, A., 1980. Rapid taste-aversion learning by an isolated molluscan central nervous system. Proceedings of the National Academy of Sciences, USA 77 (10), 6204–6206.

Chung, S. T. L., Levi, D. M., Li, R. W., 2006. Learning to identify contrast-defined letters in peripheral vision. Vision Res 46 (6-7), 1038–1047.

Cohen, J., 1988. Statistical power analysis for the behavioral sciences, 2nd Edition. L. Erlbaum Associates, Hillsdale, N.J.

Cook, R., Fagot, J., Jun 2009. First trial rewards promote 1-trial learning and prolonged memory in pigeon and baboon. Proc Natl Acad Sci U S A 106 (23), 9530–9533.

Fahle, M., Edelman, S., Poggio, T., 1995. Fast perceptual learning in hyperacuity. Vision Res 35 (21), 3003–3013.

Fiorentini, A., Berardi, N., 1981. Learning in grating waveform discrimination: specificity for orientation and spatial frequency. Vision Research 21 (7), 1149–1158.

Furmanski, C. S., Engel, S. A., 2000. Perceptual learning in object recognition: object specificity and size invariance. Vision Res 40 (5), 473–484.

Gold, J., Bennett, P. J., Sekuler, A. B., 1999a. Identification of band-pass filtered letters and faces by human and ideal observers. Vision Res 39 (21), 3537–3560.

Gold, J., Bennett, P. J., Sekuler, A. B., 1999b. Signal but not noise changes with perceptual learning. Nature 402 (6758), 176–178.

Gold, J. M., Sekuler, A. B., Bennett, P. J., 2004. Characterizing perceptual learning with external noise. Cognitive Sci 28, 167–207.

Hauptmann, B., Reinhart, E., Brandt, S. A., Karni, A., 2005. The predictive value of the leveling off of within session performance for procedural memory consolidation. Brain Res Cogn Brain Res 24 (2), 181–189.

Hawkey, D., Amitay, S., Moore, D., 2004. Early and rapid perceptual learning. Nature Neuroscience 7 (10), 1055–1056.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics 6, 65–70.

Hothorn, T., Bretz, F., Westfall, P., 2008. Simultaneous inference in general parametric models. Biometrical Journal 50 (3), 346–363.

Husk, J. S., Bennett, P. J., Sekuler, A. B., 2007. Inverting houses and textures: investigating the characteristics of learned inversion effects. Vision Res 47 (27), 3350–3359.

Hussain, Z., Bennett, P., Sekuler, A., 2005. Perceptual learning of faces and textures is tuned to trained identities. Perception: ECVP Abstract Supplement 34.

Hussain, Z., Bennett, P., Sekuler, A., 2007. Superior identification of familiar visual patterns a year after learning. J Vision 7 (9), 797a.

Hussain, Z., Sekuler, A., Bennett, P., 2009. Contrast reversal abolishes perceptual learning. Journal of Vision 9 (4), 1–8.

Hussain, Z., Sekuler, A., Bennett, P., In Press. Perceptual learning modifies inversion effects for faces and textures. Vision Research.

Karni, A., Bertini, G., 1997. Learning perceptual skills: behavioral probes into adult cortical plasticity. Curr Opin Neurobiol 7 (4), 530–535.

Karni, A., Sagi, D., 1993. The time course of learning a visual skill. Nature 365 (6443), 250–252.

Kourtzi, Z., Betts, L., Sarkheil, P., Welchman, A., 2005. Distributed neural plasticity for shape learning in the human visual cortex. PLOS Biology 3 (7).

Maertens, M., Pollmann, S., 2005. fmri reverals a common neural substrate of illusory and real contours in v1 after perceptual learning. Journal of Cognitive Neuroscience 17 (10), 1553–1564.

Malin, D., Jenkins, P., Watts, M., Spezia, P., Novy, B., 1986. One-trial discriminative reward learning- reward magnitude and spatial reversal effects. Pharmacology Biochemistry and Behavior 25 (1), 303.

Matthews, N., Liu, Z., Geesaman, B., Qian, N., 1999. Perceptual learning on orientation and direction discrimination. Vision Research 30 (22), 3692–3701.

Mednick, S. C., Arman, A. C., Boynton, G. M., 2005. The time course and specificity of perceptual deterioration. Proc Natl Acad Sci U S A 102 (10), 3881–3885.

Ofen, N., Moran, A., Sagi, D., 2007. Effects of trial repetition in texture discrimination. Vision Res 47 (8), 1094–1102.

Ofen-Noy, N., Dudai, Y., Karni, A., 2003. Skill learning in mirror reading: how repetition determines acquisition. Cognitive Brain Research 17, 507–521.

Pelli, D. G., 1997. The videotoolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10 (4), 437–442.

Poggio, T., Fahle, M., Edelman, S., 1992. Fast perceptual learning in visual hyperacuity. Science 256 (5059), 1018–21.

Pourtois, G., Rauss, K., Vuilleumier, P., Schwarz, S., 2008. Effects of perceptual learning on primary visual cortex activity in humans. Vision Research 48 (1), 55–62.

R Development Core Team, 2007. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (http://www.R-project.org).

Raiguel, S., Vogels, R., Mysore, S., Orban, G., 2006. Learning to see the difference specifically alters the most informative v4 neurons. The Journal of Neuroscience 26 (24), 6589–6602.

Rainer, G., Lee, H., Logothetis, 2004. The effect of learning on the funcion of monkey extrastriate visual cortex. PLOS Biology 3 (2).

Richards, E., Bennett, P., Sekuler, A., 2006. Age related differences in learning with the useful field of view. Vision Research 46 (25), 4217–4231.

Rozin, P., 1986. One-trial acquired likes and dislikes in humans - disgust as a us, food predominance, and negative learning predominance. Learning and Motivation 17 (2), 180–189.

Sahley, C., Gelperin, A., Rudy, J., 1981. One-trial associative learning modifies food odor preferences of a terrestrial mollusk. Proccedings of the National Academy of Sciences of the United States of America- Biological Sciences 78 (1), 640–642.

Savion-Lemieux, T., Penhune, V., 2005. The effects of practice and delay on motor skill learning and retention. Experimental Brain Research 161, 423–431.

Schoups, A., Vogels, R., Qian, N., Orban, G., 2001. Practising orientation identification improves orientation coding in v1 neurons. Nature 412, 549–553.

Schoups, A. A., Vogels, R., Orban, G. A., 1995. Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. J Physiol 483 (Pt 3), 797–810.

Schwartz, S., Maquet, P., Frith, C., 2002. Neural correlates of perceptual learning: a functional mri study of visual texture discrimination. Proc Natl Acad Sci U S A 99 (26), 17137–17142.

Sigman, M., Gilbert, C. D., 2000. Learning to find a shape. Nat Neurosci 3 (3), 264–269.

Sireteanu, R., Rettenbach, R., 1995. Perceptual learning in visual search: Fast, enduring but non-specific. Vision Research 35 (14), 2037–2043.

Sowden, P., Rose, D., Davies, I., 2002. Perceptual learning of luminance contrast detection: specific for spatial frequency and retinal location but not orientation. Vision Research 42 (10), 1249–1258.

Taieb-Maimon, M., 2007. Learning headway estimation in driving. Human Factors 49 (4), 734–744.

Tovee, M., Rolls, E., Ramachandran, V., 1996. Rapid visual learning in the neurones of the primate temporal visual cortex. Neuroreport 7 (15-17), 2757–2760.

Tyler, C. W., Liu, H. C. L., McBride, B., Kontsevich, L., 1992. Bit-stealing: How to get 1786 or more gray levels for an 8-bit color monitor. In: Rogowitz, B. E. (Ed.), Proceedings of SPIE: Human vision, visual processing and digital display III. Vol. 1666. pp. 351–364.

Westfall, P., 1997. Multiple testing of general contrasts using logical constraints and correlations. Journal of the American Statistical Association 92 (437), 299–306.

Wilcox, R. R., 2005. Introduction to robust estimation and hypothesis testing, 2nd Edition. Elsevier/Academic Press, Amsterdam.

Woodsworth, R., Schlosberg, H., 1956. Psychology, revised Edition. Henry Holt and Company, Inc.

Wright, B., Sabin, A., 2007. Perceptual learning: how much daily training is enough? Experimental Brain Research 180 (4), 727–736.

Yi, D.-J., Olson, I. R., Chun, M. M., 2006. Shape-specific perceptual learning in a figure-ground segregation task. Vision Res 46 (6-7), 914–924.

Yotsumoto, Y., Watanabe, T., Sasaki, Y., 2008. Different dynamics of performance and brain activation in the time course of perceptual learning. Neuron 57 (6), 827–833.

Chapter 6

Superior identification of familiar visual patterns a year after learning

6.1 Abstract

We examined whether learning in face and texture identification tasks is stable across long periods of time. Subjects completed 10AFC identification tasks on two consecutive days and returned for a follow-up session 10-18 months later. Response accuracy improved significantly across the first two days of training, and most of the benefits of practice remained in the follow-up session approximately one year later. The learning and retention was largely stimulus-specific: novel exemplars from the trained category did not exhibit all the benefits found with the trained set. Individual differences in performance observed during the first two days were preserved at the follow-up session. The relative difficulty of identifying individual stimuli was stable across testing sessions. These results are similar to the stability of learning found with low-level visual tasks. We conclude that perceptual learning of complex visual tasks endures over time.

6.2 Introduction

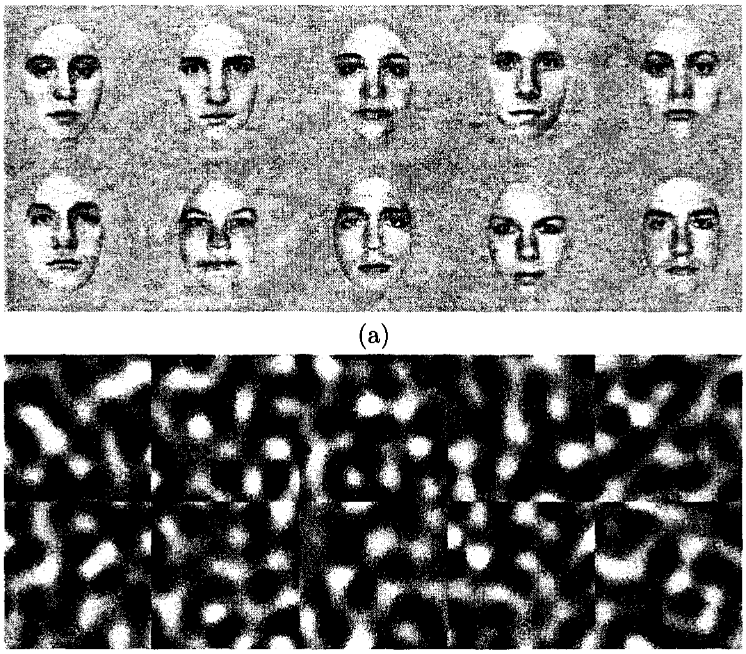
It is well known that practice improves visual performance on simple tasks such as the discrimination of spatial frequency, orientation, direction of motion and curvature (Fiorentini and Berardi, 1981; Karni and Sagi, 1993; Ball and Sekuler, 1982; Maertens and Pollmann, 2005). Practice also improves contrast sensitivity in amblyopes (Zhou et al., 2006), and grating acuity and figure-ground segregation in typical observers (Beard et al., 1995; Yi et al., 2006). Improvements on these tasks are specific to the trained stimuli (Ball and Sekuler, 1982; Fiorentini and Berardi, 1981; Yi et al., 2006), and endure over long periods of time without intervening training (Fiorentini and Berardi, 1981; Beard et al., 1995; Ball and Sekuler, 1982; Karni and Sagi, 1993; Zhou et al., 2006; Yi et al., 2006; Maertens and Pollmann, 2005). These specific, enduring improvements with simple tasks suggest the permanent alteration of early visual areas where the stimuli used are sufficiently represented (i.e., areas V1-V4). Practice also improves performance in more complex visual tasks. For example, practice improves reading of inverted text, visual search, and face- and texture identification (Kolers, 1976; Sigman and Gilbert, 2000; Sireteanu and Rettenbach, 1995; Gold et al., 1999b, 2004). As with simple tasks, the benefits of practice in complex tasks can be stimulus-specific (Kolers, 1976; Sigman and Gilbert, 2000; Hussain et al., 2005), suggesting that learning in complex situations shares some of the characteristics of learning found in reduced contexts.

Do the stimulus-specific effects of training persist for complex tasks? In one striking example, subjects reread inverted text encountered a year earlier more rapidly than novel pages of inverted text (Kolers, 1976), indicating a selective, persisting improvement of the encoding operations involved. We have obtained similar specificity of learning in the short-term, with face- and texture identification: improvements in identifying faces and textures are largely confined to the trained items; at test, trained items are identified at an advantage relative to novel items (Hussain et al., 2005, In Press). Here, we test whether this type of stimulus-specificity endures. It could be that the learning is entirely transient and all improvements vanish with time, or some generalized benefits might be retained, such as those found in learning of visual search (Sireteanu and Rettenbach, 1995), and useful field of view (Richards et al., 2006). Alternately, learning of faces and textures could resemble that found with simple tasks, with stimulus-specific benefits persisting up to a year later.

6.3 Methods

6.3.1 Subjects

Nine subjects performed a face-identification task, and six subjects performed a texture-identification task on two occasions: an initial learning and test phase in experiments that were conducted with a larger sample of observers (168 and 158, for faces and textures, respectively), and a follow-up test phase 10-18 months after initial training (henceforth referred to as "1 year later"). This subset of observers was selected not on the



(a)

(b)

Figure 6.1: Examples of the faces (a) and textures (b) used in the experiments.

basis of any performance criterion, but on availability to return at the time the follow-up study was conducted. On average, subjects in the face- and texture identification task performed the follow-up session respectively 13 (SD = 3.4) and 15 months (SD =1.6) after the initial test session. Subjects were between the ages of 17 and 32 years (M = 19.26 SD = 2.18) and had normal or corrected-to-normal Snellen acuity. All subjects received a small fee ($10/hour) or partial course credit for participating in the experiment.

6.3.2 Apparatus and Stimuli

Stimuli were generated in Matlab (v. 5.2) using the Psychophysics and Video Toolboxes (Brainard, 1997; Pelli, 1997), and displayed on a 21" Sony Trinitron monitor (1024 x 768 pixels) at a frame rate of 85 Hz. Average luminance was 62 cd/m^2. Display luminance was measured with a PhotoResearch PR650 photometer, and the calibration data were used to build a 1779-element lookup table (Tyler et al., 1992). Customized computer software constructed the stimuli on each trial by selecting the appropriate luminance values from the calibrated lookup table and storing them in the display's eight-bit lookup table.

Figure 6.1 shows examples of the face and texture stimuli presented to subjects. The methods used to create face stimuli have been described previously (Gold et al., 1999a,b) and therefore only the main points are listed here. Faces were cropped to show only internal features and equated in terms of their amplitude spectrum. The textures were band-limited noise patterns created by applying an isotropic, ideal band-pass spatial frequency filter (2-4 cy/image) to white Gaussian noise. Stimulus size for both stimulus classes was 256 x 256 pixels, subtending 4.8 x 4.8 deg of visual angle from the viewing distance of 114 cm. Two sets of 10 faces (five female and five male in each set) and two sets of 10 textures were created. During the experiment, stimuli were presented in low, medium, or high levels of static two-dimensional Gaussian noise (contrast variance = 0.001, 0.01, or 0.1), and stimulus contrast was varied across trials using the method of constant stimuli. Seven levels of contrast that were spaced equally on a logarithmic scale were chosen at each noise level, such that the contrasts spanned a range that was sufficient to produce significant changes in performance in virtually all subjects. The combination of stimulus contrast levels and noise levels yielded 21 different stimulus presentation conditions. Hence, subjects viewed each face/texture at a signal-to-noise ratio that varied significantly across trials.

6.3.3 Procedure

All subjects performed two sessions of the face- or texture identification task at the same time on consecutive days (Days 1 and 2). Observers then were tested in a final session approximately one year later with two sets of stimuli: one that contained the same stimuli viewed in the first two sessions, and another that contained items that they had not not seen previously. The same and different sets of stimuli were presented in separate blocks of trials, and the order was roughly counterbalanced across subjects.

Subjects were seated in a darkened room 114 cm away from the monitor. Viewing was binocular, and viewing position and distance were stabilized with an adjustable chinrest. The experiment started after a 60 s period during which the subject adapted to the average luminance of the display. A trial began with the presentation of a black, high-contrast fixation point (0.15 x 0.15 deg) in the center of the screen for 100 ms, followed by a face/texture, selected randomly from one of the 21 stimulus conditions, presented for 200 ms at the center of the screen. After the face/texture disappeared, the entire set of 10 faces/textures was presented as noiseless, high-contrast thumbnail images, each subtending 1.7 x 1.7 deg of visual angle. Five thumbnails were presented on the top half of the screen, and five on the bottom half, and the location of each face/texture in the response window was constant across trials and days. The subject's task was to decide which one of the 10 faces/textures had been presented during the trial by clicking on the chosen texture with a computer mouse. Auditory feedback in the form of high-pitched (correct) and low-pitched (incorrect) tones informed the subject about the accuracy of each response, and the next trial began one second after presentation of the feedback.

On each of Days 1 and 2, observers performed 40 trials per stimulus condition for a total of 840 trials per day (40 trials x 21 stimulus conditions), which were completed in approximately one hour. Stimulus sets remained the same for an individual across Days 1 and 2. Each face/texture was selected randomly (with replacement) on each trial, such that on average each face/texture was shown approximately 84 times during the entire session. The follow-up session (1-year later) consisted of 420 trials with the trained stimulus set, and 420 trials with the novel stimulus set for the same stimulus class (faces or textures). The trained and novel stimuli were shown in separate blocks in an order roughly counterbalanced across subjects (full counterbalanced for the six subjects in the texture identification task, and the almost fully counterbalanced for the nine subject in the face task).

6.4 Results

For the purpose of the analyses, the 840 trials on Days 1 and 2 were divided into four blocks of 210 sequential trials (Trial bins 1-8), and the 420 trials for each of the stimulus sets tested a year later were divided into two blocks of 210 trials (Bins A and B). For each bin, the proportion of correct responses was calculated after collapsing across all levels of stimulus contrasts and noise. Results for face- and texture-identification are shown in Figures 6.2 and 6.3.

The face identification task showed significant within-session improvement on Day 1, evidenced by a 20% increase in accuracy from Bin 1 to Bin 4 ($t(8) = 9.54$, $p < .00001$). There was also significant within-session improvement on Day 2, with performance improving by 10% from Bin 5 to Bin 8 ($t(8) = 4.54$, $p < .01$). Overall accuracy – i.e., averaged across bins – was 18% higher on Day 2 than on Day 1, ($t(8) = 11.36$, $p < .0001$). A year later, performance in the first trial bin (Bin A) in the same-face condition was 18% better than in Bin 1 ($t(8) = 7.883$, $p < .0001$), no different than performance in Bin 4 ($t(8) = 1.054$, $p = .322$), but 15% lower than performance in Bin 8 ($t(8) = 7.17$, $p < .0001$). Therefore, performance in the follow-up session in the same-face condition was equivalent to that achieved at the end of the first, but not the second, session. In the different-face condition, performance in Bin A was no different than performance

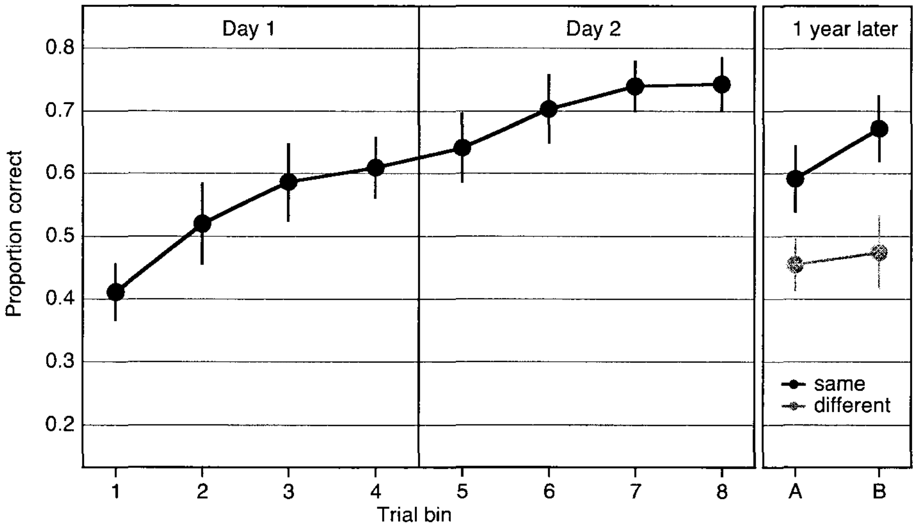


Figure 6.2: Proportion correct on Day 1, Day 2, and 1 year later for 9 subjects in the face-identification task. Each session comprised 840 trials, therefore each trial bin represents 210 trials. Performance a year later was measured both with the same faces shown on Days 1 and 2 (black circles), and with different faces that subjects had not seen before (gray circles).

in Bin 1 ($t(8) = 1.8, p = .10$) and 15% lower than performance in Bin 4 ($t(8) = 4.05$, $p = .0036$). Additionally, performance in the same-face condition was better than performance in the different-face condition in Bin A (14%; $t(8) = 3.57$, $p = .007$) and Bin B (20%; $t(8) = 4.81$, $p = .001$). A two-way ANOVA on the follow-up data, with Bin (A vs. B), and Condition (Same vs. Different), as repeated measures, yielded a main effect of Condition ($F(1,8) = 23.86, p = .001$), and a main effect of Bin ($F(1,8) = 6.11, p = .038$); the interaction between Bin and Condition was not significant ($F(1,8) = 2.37, p = .161$). This analysis confirms that accuracy in the follow-up session was higher in the same-face than in the different-face condition, and that accuracy increased from Bin A to Bin B; however the amount of improvement from Bin A to Bin B did not differ across conditions.

The texture-identification task also showed significant within-session learning on Day 1: accuracy increased by 25% from Bin 1 to Bin 4 ($t(5) = 5.070$, $p = .003$). There also was significant within-session learning on Day 2, by 16% from Bin 5 to Bin 8 ($t(5) = 2.96$, $p = .03$). Averaged across bins, response accuracy on Day 2 was 21% higher than on Day 1 ($t(5) = 8.18$, $p = .0004$). Approximately one year later, accuracy in Bin A in the same-texture condition was 35% higher than in Bin 1 ($t(5) = 15.70$, $p < .0001$) and 9% higher than in Bin 4 ($t(5) = 2.68$, $p = .04$), and did not differ significantly from

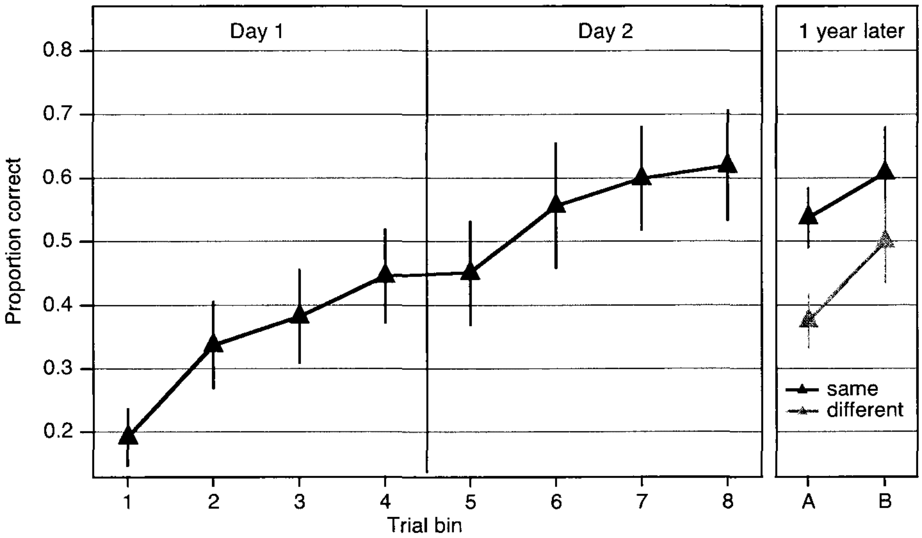


Figure 6.3: Proportion correct on Day 1, Day 2 and 1 year later for 6 subjects in the texture-identification task. Each session comprised 840 trials, therefore each trial bin represents 210 trials. Performance a year later was measured both with the same faces shown on Days 1 and 2 (black triangles), and with different faces that subjects had not seen before (gray triangles).

accuracy in Bin 8 ($t(5) = 1.54$, $p = .18$). In the different-texture condition, accuracy in Bin A was 18% better than in Bin 1 ($t(5) = 6.41$, $p = .001$), no different than in Bin 4 ($t(5) = 1.87$, $p = .11$), and 24% lower than in Bin 8 ($t(5) = 4.88$, $p = .004$). Therefore, there was partial generalization to novel items, but performance was not as good as that achieved with the same items. Performance in the same-texture condition was better than in the different-texture condition in Bin A (16% difference; $t(5) = 6.44$, $p = .0013$) and Bin B (10% difference; $t(5) = 3.29, p = .02$), indicating that a substantial proportion of the improvements was stimulus-specific. A two-way ANOVA on the follow-up data, with Bin (A vs. B), and Condition (Same vs. Different), as repeated measures, yielded a main effect of Condition ($F(1,5) = 34.32, p = .002$), and a main effect of Bin ($F(1,5) = 14.103, p = .01$); the interaction between Bin and Condition was not significant ($F(1,5) = 2.17, p = .20$). This analysis confirms that as with faces, accuracy was higher in the same-texture than in the different-texture condition, accuracy increased from Bin A to Bin B, and the amount of improvement from Bin A to Bin B did not differ across conditions.

To test whether the amount of transfer to novel items differed for faces and textures, the difference between Bin 1 and Bin A was compared in the different-texture and different-face conditions. There was a significant difference between faces and textures in the amount of transfer to novel items (faces: 4% versus textures: 18%; $t(10.873) = 3.79, p = .003$). However, the advantage of same over different items in Bin A did not differ between faces and textures (faces: 14% versus textures: 16%; $t(12.637) = .55, p = .58$), and neither did the advantage of same over different items in Bin B (faces: 19% versus textures: 11%, $t(12.998) = 1.69, p = .11$). These analyses show that there was more generalization to novel items with the texture stimuli, nevertheless the stimulus-specific advantage did not differ between faces and textures.

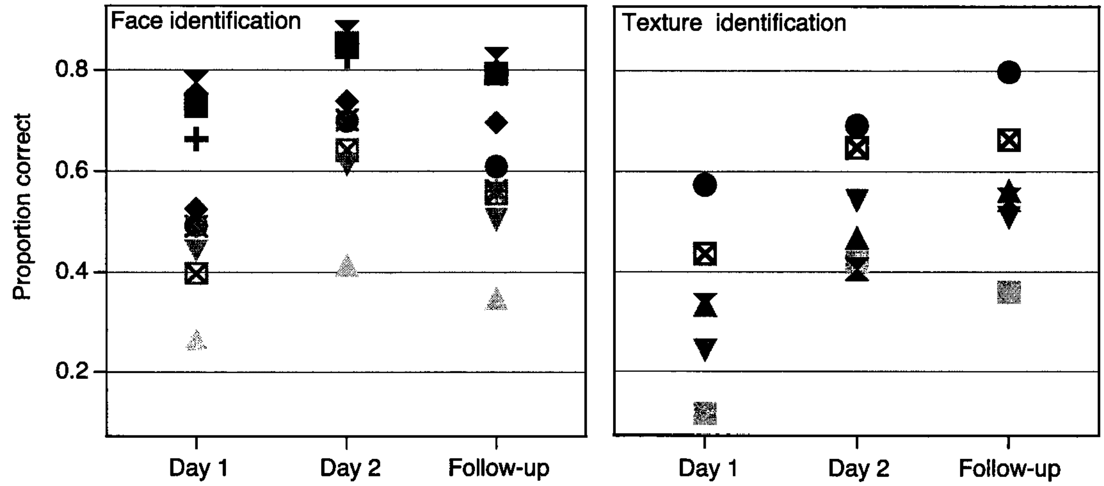


Figure 6.4: Individual performance on the face and texture identification tasks on Day 1, Day 2, and during the follow-up session about one year later. Each symbol represents a subject.

Figure 6.4 shows the average performance of each observer on each session. Despite the large variability in performance across observers, the rank orders of individuals was relatively stable across sessions in both tasks. For subjects who were tested with faces, the Spearman rank-order correlation was significant for all pairs of test sessions ($\rho \geq .96$, $p < .001$). For subjects who were tested with textures, the rank-order correlations were all positive, but, perhaps due to the small sample size ($n = 6$), only the correlation between Day 1 and the follow-up session was significant (Days 1 & 2: $\rho = 0.60$, $p = .24$; Day 1 & follow-up: $\rho = .94$, $p = .015$; Day 2 & follow-up: $\rho = .71$, $p = .13$).

To examine whether the effects of practice reflected improved performance on just a few items, we calculated the mean response accuracy for individual faces and textures during each test session. The results for faces are shown in Figure 6.5. Each symbol in

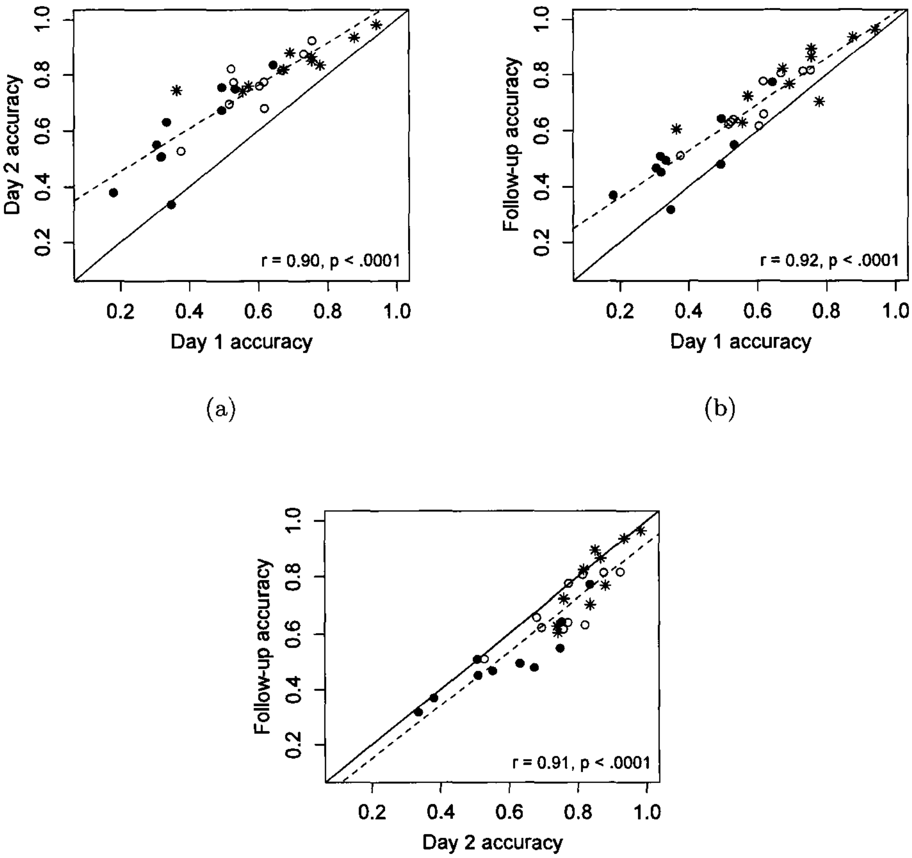
Figure 6.5 represents the response accuracy for a single face averaged across subjects. Three sets of ten faces are shown by different symbols, because average accuracy for each set was calculated with different numbers of subjects; the face sets shared the same stimulus characteristics, but were used in different experiments. Figures 6.5a and 6.5b show that relative to Day 1, accuracy was higher for 29 out of 30 faces on Day 2, and 27 out of 30 faces during follow-up. Accuracy during follow-up was no different than accuracy measured on Day 2 (Figure 6.5c), confirming that the benefits of learning were intact during the follow-up session. The variation among all faces remained constant across all sessions (all r's $\geq .90$, $p < .001$), implying that learning did not change the relative discriminability of the faces.

Figure 6.6 shows the effects of practice on response accuracy for individual textures. As in Figure 6.5, each symbol represents the accuracy (averaged across subjects) for a single texture. Compared to Day 1, response accuracy was higher for every texture on Day 2 (Figure 6.6a) and for 19 out of 20 textures on the 1-year follow-up (Figure 6.6b). Accuracy during follow-up was no different than that measured on Day 2. Hence, practice improved performance for all textures, and the benefits relative to Day 1 were retained for nearly all items. Finally, the variation among textures remained stable across all sessions (all r's $\geq .68$, $p < .001$). This result is consistent with the idea that learning (or forgetting) did not produce qualitative changes in the way these items were represented.

6.5 Discussion

Even for complex visual tasks such as face- and texture identification, the effects of perceptual learning were remarkably stable over long periods of time. The lasting effects were largely specific to the trained stimuli (although there was some generalization with the textures), similar to the specificity of learning observed a day after training with faces and textures (Hussain et al., 2005, In Press, 2009). In addition to the stimulus-specificity of learning, individual differences from the initial sessions were preserved a year later, and so were inter-item differences in accuracy. Item-accuracy was positively correlated across days, and during the follow-up session, indicating that training did not alter the relative discriminability of the stimuli i.e., enhance the representation of a subset of the exposed items. Instead, learning resulted in a uniform increase in performance for all trained items. This result constrains models of learning that alter the stimulus space after training.

It is striking that the enduring benefits for these stimuli were at all item-specific.

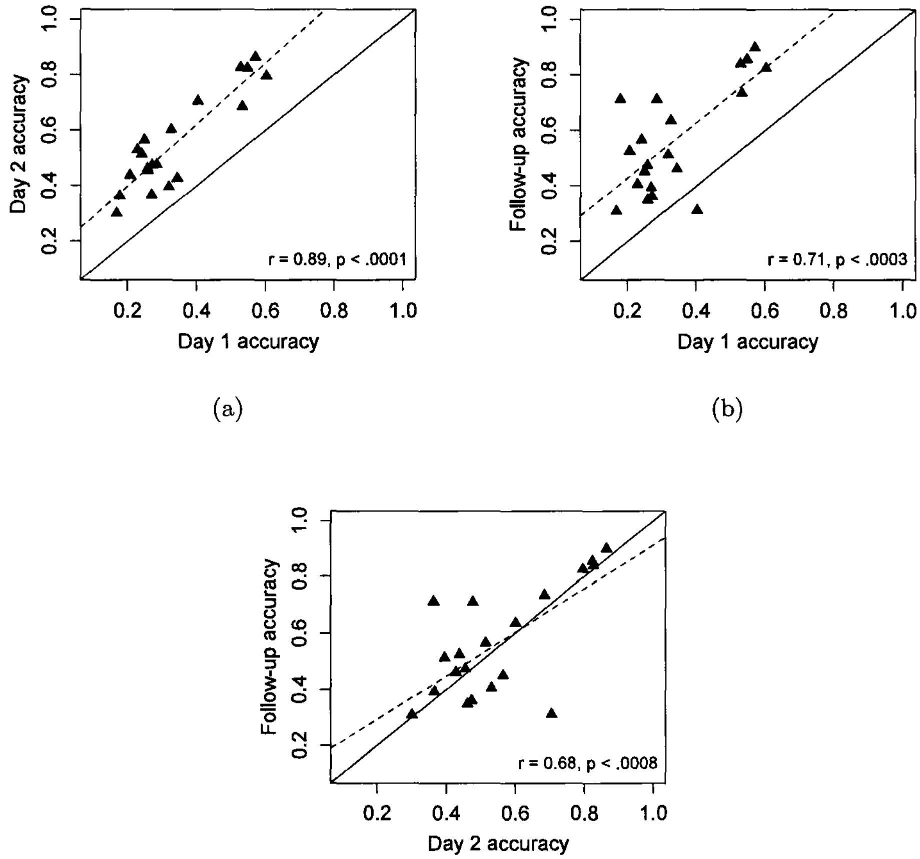


(a)

(b)

(c)

Figure 6.5: Scatter plots showing accuracy for three sets of ten faces used in the face identification task; the sets were used in separate experiments, and are shown by separate symbols. Black circles, open circles and stars respectively represent the average performance of four, two and three subjects with a given face. Area above the solid line indicates improvement. Dashed line indicates the least squares fit. a) Day 1 vs. Day 2. b) Day 1 vs. 1 year later c) Day 2 vs. 1 year later. All correlations are positive, and significant. Learning is retained a year later relative to performance on Day 1 (b), however there is a drop relative to performance on Day 2 (c).



(a)

(b)

(c)

Figure 6.6: Scatter plots showing accuracy for each of twenty textures used in the texture identification task. Each point represents the average performance of three subjects with a given texture. Area above the solid line indicates improvement. Dashed line indicates the least squares fit. a) Day 1 vs. Day 2. b) Day 1 vs. 1 year later c) Day 2 vs. 1 year later. All correlations are positive, and significant. Learning is retained a year later relative to performance on Day 1 (b), and some improvement relative to performance on Day 2 is also evident (c).

The textures are unfamiliar stimuli without the spatial characteristics found in everyday objects, and without any obvious semantic content. Nevertheless, performance with trained items a year later was superior to novel items, indicating that familiarity with the object class is not essential for long-term retention. The enduring specificity of learning found with faces is also notable given the large amount of exposure to other faces that presumably occurred for all participants in natural contexts during the intervening period; Evidently, exposure to other faces does not dilute the specific effects of perceptual learning up to a year later.

Textures evidenced more generalization of learning to untrained items than did faces. Specifically, performance with different textures (but not faces) in the 1-year follow-up was superior to performance measured in the first trial bin on Day 1. Previously we have found small amounts of generalization to novel items both with faces and with textures, and attributed the effect to a task-general component of learning (Hussain et al., In Press), consistent with the mixture of specificity and transfer reported in other perceptual learning studies (Yi et al., 2006; Beard et al., 1995). Presumably, the larger amounts of transfer found with textures is due to subjects becoming accustomed in the initial trial bin to the arbitrary quality of the texture patterns.

Lasting, stimulus-specific effects of learning on simple visual discriminations have been attributed to changes early in the visual pathway, particularly, to the primary visual area (Karni and Bertini, 1997; Schoups et al., 1995; Maertens and Pollmann, 2005). Clearly, similar learning is possible for high-level visual tasks that in all likelihood, engage brain regions beyond the primary visual area. Complex objects composed of multiple stimulus features, such as faces and textures, are coded in entirety by neurons in higher visual areas such as inferotemporal cortex (IT) (Desimone et al., 1984; Logothetis et al., 1995). Neurons in IT are highly stimulus selective, yet the selectivity of these neurons can be profoundly affected by learning (Cox and DiCarlo, 2008; Li and DiCarlo, 2008). IT neurons are also involved in the formation of visual memory (Desimone, 1996). Stimulus-selective coding of objects could be an outcome of repeated exposure, and a factor in the stable effects of training that we show here for faces and textures.

Although one might infer that long-lasting perceptual improvements are the result of neural changes that are irrevocable, a recent report suggests otherwise: increased activation in V1 was associated with stimulus-specific improvements on a texture discrimination task, but performance remained stable while activation in the corresponding regions returned to baseline (Yotsumoto et al., 2008). This finding implies the involvement of additional neural regions in enabling the long-term effects of learning, besides the neurons that code immediate properties of the trained stimuli. It could be that task context activates a larger neural network that re-instantiates the effects of learning long after the initial training experience.

6.6 Conclusions

Perceptual learning of face- and texture-identification endures over long periods of time, here measured as 10-18 months after initial training. The retained improvements were largely specific to the trained items, and thus cannot be considered as the retention of task-general skills, although some generalized benefits were also evident during follow-up, particularly for texture-identification. Individual differences, as well as inter-item differences in accuracy were intact during follow-up, indicating that learning did not qualitatively alter the representations of the learned items. The long-term, specific effects of perceptual learning for complex patterns suggest similar principles of stimulus coding in higher visual areas as in earlier visual areas typically associated with the learning of reduced stimuli.

References

Ball, K., Sekuler, R., 1982. A specific and enduring improvement in visual-motion discrimination. Science 218 (4573), 697–698.

Beard, B., Levi, D., Reich, L., 1995. Perceptual learning in parafoveal vision. Vision Research 35 (12), 914–924.

Brainard, D. H., 1997. The psychophysics toolbox. Spat Vis 10 (4), 433–436.

Cox, D. D., DiCarlo, James, J., October 2008. Does learned shape selectivity in inferior temporal cortex automatically generalize across retinal position? The Journal of Neuroscience 28 (40), 10045–10055.

Desimone, R., November 1996. Neural mechanisms for visual memory and their role in attention. Proceedings of the National Academy of Sciences, USA 93, 13494–13499.

Desimone, R., Albright, T., Gross, C., Bruce, C., 1984. Stimulus-selective properties of inferior temporal neurons in the macaque. The Journal of Neuroscience 4 (8), 2051–2062.

Fiorentini, A., Berardi, N., 1981. Learning in grating waveform discrimination: specificity for orientation and spatial frequency. Vision Research 21 (7), 1149–1158.

Gold, J., Bennett, P. J., Sekuler, A. B., 1999a. Identification of band-pass filtered letters and faces by human and ideal observers. Vision Res 39 (21), 3537–3560.

Gold, J., Bennett, P. J., Sekuler, A. B., 1999b. Signal but not noise changes with perceptual learning. Nature 402 (6758), 176–178.

Gold, J. M., Sekuler, A. B., Bennett, P. J., 2004. Characterizing perceptual learning with external noise. Cognitive Sci 28, 167–207.

Hussain, Z., Bennett, P., Sekuler, A., 2005. Perceptual learning of faces and textures is tuned to trained identities. Perception: ECVP Abstract Supplement 34.

Hussain, Z., Sekuler, A., Bennett, P., 2009. Contrast reversal abolishes perceptual learning. Journal of Vision 9 (4), 1–8.

Hussain, Z., Sekuler, A., Bennett, P., In Press. Perceptual learning modifies inversion effects for faces and textures. Vision Research.

Karni, A., Bertini, G., 1997. Learning perceptual skills: behavioral probes into adult cortical plasticity. Curr Opin Neurobiol 7 (4), 530–535.

Karni, A., Sagi, D., 1993. The time course of learning a visual skill. Nature 365 (6443), 250–252.

Kolers, P., 1976. Pattern-analyzing memory. Science 191 (4233), 1280–1281.

Li, N., DiCarlo, J., September 2008. Unsupervised natural experience rapidly alters invariant object representation in visual cortex. Science 321, 1502–1506.

Logothetis, N., Pauls, J., Poggio, T., 1995. Shape representation in the inferior temporal cortex of monkeys. Current Biology 5 (5), 552–563.

Maertens, M., Pollmann, S., 2005. fmri reverals a common neural substrate of illusory and real contours in v1 after perceptual learning. Journal of Cognitive Neuroscience 17 (10), 1553–1564.

Pelli, D. G., 1997. The videotoolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10 (4), 437–442.

Richards, E., Bennett, P., Sekuler, A., 2006. Age related differences in learning with the useful field of view. Vision Research 46 (25), 4217–4231.

Schoups, A. A., Vogels, R., Orban, G. A., 1995. Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. J Physiol 483 (Pt 3), 797–810.

Sigman, M., Gilbert, C. D., 2000. Learning to find a shape. Nat Neurosci 3 (3), 264–269.

Sireteanu, R., Rettenbach, R., 1995. Perceptual learning in visual search: Fast, enduring but non-specific. Vision Research 35 (14), 2037–2043.

Tyler, C. W., Liu, H. C. L., McBride, B., Kontsevich, L., 1992. Bit-stealing: How to get 1786 or more gray levels for an 8-bit color monitor. In: Rogowitz, B. E. (Ed.), Proceedings of SPIE: Human vision, visual processing and digital display III. Vol. 1666. pp. 351–364.

Yi, D.-J., Olson, I. R., Chun, M. M., 2006. Shape-specific perceptual learning in a figure-ground segregation task. Vision Res 46 (6-7), 914–924.

Yotsumoto, Y., Watanabe, T., Sasaki, Y., 2008. Different dynamics of performance and brain activation in the time course of perceptual learning. Neuron 57 (6), 827–833.

Zhou, Y., Huang, C., Xu, P., Tao, L., Qiu, Z., Lu, Z., 2006. Perceptual learning improves contrast sensitivity and visual acuity in adults with anisometropic amblyopia. Vision Research 46, 739–750.

Chapter 7

Summary & the Future

Perceptual learning of faces and textures shows the hallmarks of learning found with simple visual tasks. The hallmarks are stimulus-specificity, a time-course comprising rapid and gradual improvements, and long-lasting stability of learning. Similarities in learning of simple and complex visual patterns suggest common principles of learning throughout the hierarchy of visual processing. Overall, the experiments show that experience elicits significant and stable improvements in the quality of complex visual representations.

7.0.1 Specificity

Stimulus-specific learning permeates aspects of object recognition thought to be inherent, such as the face-inversion effect. Chapter 2 showed that *orientation-specific* learning contributes significantly to the inversion effect, which either increases or decreases, depending on the orientation of the trained faces. Furthermore, practice generates inversion effects where none exist: in textures with no canonical orientation, and without the prototypical structure present in faces. Clearly, the face inversion-effect is not entirely inherent, and other supposedly inherent effects could be labile. Chapter 2 also identifies learning as largely *exemplar-specific*, which makes perceptual learning a likely mechanism for phenomena in which certain visual exemplars from an object class are distinguished better than others - for instance, the 'other-race effect', where own-race faces are recognized better than other-race faces (Valentine and Bruce, 1986; Goldstein and Chance, 1985). Stimulus-specific perceptual learning of face- and texture identification is consistent with the view that the 'special effects' found with faces are largely the product of expertise after a lifetime of exposure (Gauthier et al., 1998; Bukach et al., 2006).

Chapter 3 showed that perceptual learning of textures was *contrast polarity-specific*. Learning was abolished when textures were reversed in polarity, even though the shape and location of the telling features within the image had not been altered. Learning was abolished even when observers were shown the type of change the stimuli had undergone, and when they were given the opportunity to modify their learning strategies to anticipate the change. These experiments provide the most vivid demonstration of specificity of learning to the exact properties of the stimuli, and they show how stimulus-driven effects on learning can be robust to cognitive, 'top-down' control.

Specificity of learning has been used to infer where in the brain learning occurs (Karni and Bertini, 1997). In simple cases, stimulus- and position-specific learning is associated with changes in the primary visual area, which is retinotopically organized, and which fully represents the relevant stimulus features (Karni and Sagi, 1991; Gilbert, 1994; Sigman and Gilbert, 2000; Fahle, 2005; Yi et al., 2006). In line with this logic, one interpretation of stimulus-specific learning of faces and textures might be that practice alters the most initial encoding of the stimuli, in primary visual cortex. However, the features of complex patterns, as well as entire patterns are well-represented in higher visual areas such as inferotemporal cortex (IT) (Desimone et al., 1984; Tanaka et al., 1991; Logothetis et al., 1995). Learning of complex patterns affects the responses of IT neurons (Logothetis et al., 1995; Jagadeesh et al., 2001), and the effects have been shown as item- and location-specific (Cox and DiCarlo, 2008; Li and DiCarlo, 2008). Also, polarity-sensitive cells are profuse in IT (Ito et al., 1994). For these reasons, an equally likely interpretation of the stimulus-specific effects of learning found with faces and textures, is that the substrate of learning is higher in the visual stream. Generally, stimulus-specificity should be taken to indicate early neural changes only when the stimuli are highly reduced and the task is confined to a single stimulus dimension. In all other cases, the substrate of learning must be confirmed physiologically.

It is worth noting that some recent studies challenge the traditional inference that specificity of learning reflects modifications in primary visual areas even in highly reduced cases. Law and Gold (2008) reported significant changes in macaque area LIP with improvements on a motion discrimination task, and negligible changes in area MT; whereas area MT encodes motion, area LIP enables decision-making, suggesting that learning involves better *sampling* of the relevant encoding neurons, rather than better encoding itself. This report is consistent with the alternate, but less predominant suggestion in the literature that stimulus-specificity could correspond with a central rather than early neural substrate (Mollon and Danilova, 1996). Intriguingly, stimulus-specific perceptual learning in the auditory domain, of a challenging and complex linguistic task has recently been linked to changes outside the neocortex - in the auditory brainstem (Song et al., 2008) - raising the possibility that the neural substrate of other types of perceptual learning might also not be confined to the neocortex.

7.0.2 Time-course

Face- and texture identification are in the category of tasks showing substantial within-session effects of practice (Fiorentini and Berardi, 1981; Poggio et al., 1992; Beard et al., 1995; Matthews et al., 1999). Certain tasks do not exhibit within-session learning, and the improvements measured across sessions are attributed to consolidation that occurs during sleep (Schoups et al., 1995; Mednick et al., 2005). Chapter 4 showed that sleep had almost no impact on learning of face identification: within-session improvements were robust in the absence of sleep, and there were significant across-session gains for groups that did not sleep between sessions. These findings suggest that rather than being the outcome of a latent period of consolidation as has sometimes been proposed (Karni and Bertini, 1997), specificity of learning may emerge during training. It is not clear why within-session learning is negligible for certain tasks, and sleep's role in perceptual learning remains to be shown for the remaining variety of cases where across-session improvements have been measured.

The total amount learned on these tasks is a cumulative function of practice: for both face- and texture identification, more practice resulted in better performance, and the best performance was obtained with large amounts of practice (Chapter 5). Nevertheless, there were measurable benefits from small amounts of practice - one trial per condition in the case of face identification. The effects of small amounts of practice were evident only early in the session, because within-session learning enabled the unpracticed groups to catch up with the practiced groups. These data, as well as data from Chapter 2, show how the subtle effects of practice can be detected when performance is analyzed at a fine scale. In Chapter 2, average accuracy across trial bins obscured transfer of learning that was obvious when the early bins were inspected; in Chapter 5, average accuracy concealed the effects of small amounts of practice received the previous day. Whether the subtle effects of practice really matter when they're washed out in the average is another question, but the subtle effects are there. Overall, the amount learned in the tasks studied here was dictated by the sheer amount of practice, rather than the distribution of practice across days, or some optimal amount of practice within a day.

7.0.3 Stability

Remarkably, practice on the 10AFC task leaves permanent impressions of the images on the brain (permanent up to 18 months later, anyhow). Chapter 6 showed long-lasting effects of practice that were exemplar-specific: observers performed better with only *those* ten faces or textures that they saw months earlier, so it was not mere experience with the task that was recalled. Long-term, exemplar-specific retention of faces and textures is all the more uncanny because of the tremendous intervening exposure that subjects presumably had to other faces, and because of the meaningless, arbitrary quality of the texture patterns. One way to interpret these results is that the neural substrates of perceptual learning and long-term visual memory are the same. Indeed, IT neurons, speculated above as the substrate of learning for these tasks, show item-specific learning, and are thought to be involved in at least short-term visual memory (Desimone, 1996).

With both faces and textures, the relative discriminability of items was unchanged after learning. Accuracy increased uniformly for all trained items, rather than for a select number of items, as shown by the high correlations across days in Chapter 3 (Figure 3.4), and across months in Chapter 6 (Figures 6.5 and 6.6). The high correlations imply that the perceptual space was qualitatively unaltered after learning; differences in discriminability were highly reliable even when there was a large interval between measurements. Models of learning describe how practice refines the initial representation to have higher fidelity to the actual stimulus (Saksida, 1999; McLaren and Mackintosh, 2000). One consequence of the refinement of representations with learning, is the separation of initially clustered representations (e.g., Saksida, 1999). Where models of learning predict that practice teases apart similar representations, the present data add that given a roughly equal number of exposures for each item, practice *uniformly* increases the distance between item representations. Note however, that we did not manipulate the perceptual similarity of the faces or textures, and similarity is known to influence both the amount of learning and generalization between items (e.g., McLaren and Mackintosh, 2002). Future experiments could confirm whether the perceptual similarity of faces (or textures) affects the relative degree of improvement across items.

7.0.4 Faces versus textures

The preceding paragraphs summarized the similarities in learning of faces and textures: stimulus-specificity, some task-related transfer, long-term retention, a combination of within- and across session learning, and uniform increases in accuracy for all items. There were also some differences in learning between the two types of stimuli. For one, there was partial transfer of learning across orientation with faces, but not with textures (Chapter 5). Another difference: one practice trial per condition was enough to improve performance with faces, but not with textures. Two immediate stimulus-related factors might account for these differences: i) familiarity and ii) image properties. Faces are a highly familiar object class, whereas the textures are novel patterns. Knowing where to look in the image could speed up the learning process, and facilitate transfer to untrained views. Faces also have a clear canonical orientation, which in part is due to their prototypical structure (i.e., the average of many faces looks like a face); textures lack this property. Therefore, in addition to familiarity, the spatial regularities in faces might promote transfer of learning to other orientations. Future experiments might separate the influences of object familiarity and spatial structure on learning, and examine whether transfer of learning across orientation increases with the degree of spatial regularity within the object.

More transfer of learning across orientation with faces but not with textures is consistent with the proposal that task difficulty predicts transfer of learning, with more transfer under easy than difficult conditions (Ahissar and Hochstein, 1997; Liu and Weinshall, 2000; McLaren and Mackintosh, 2002). Texture identification was more difficult than face identification: accuracy in the initial trial bins was consistently lower on Day 1. The increased difficulty of texture identification relative to face identification was presumably due to the aforementioned differences in similarity and image properties, but this needs to be confirmed.

7.0.5 The future

An issue that has not been discussed is how to improve the learning itself. Better learning could mean faster learning, more efficient learning, or more generalization of learning. To the extent that generalization of learning is 'better', we can fiddle with the paradigm to see which one produces the best results. An intuitive prediction is that greater variation during practice will increase generalization to novel exemplars. One way to test this idea is by manipulating the standard deviation of a fixed number of

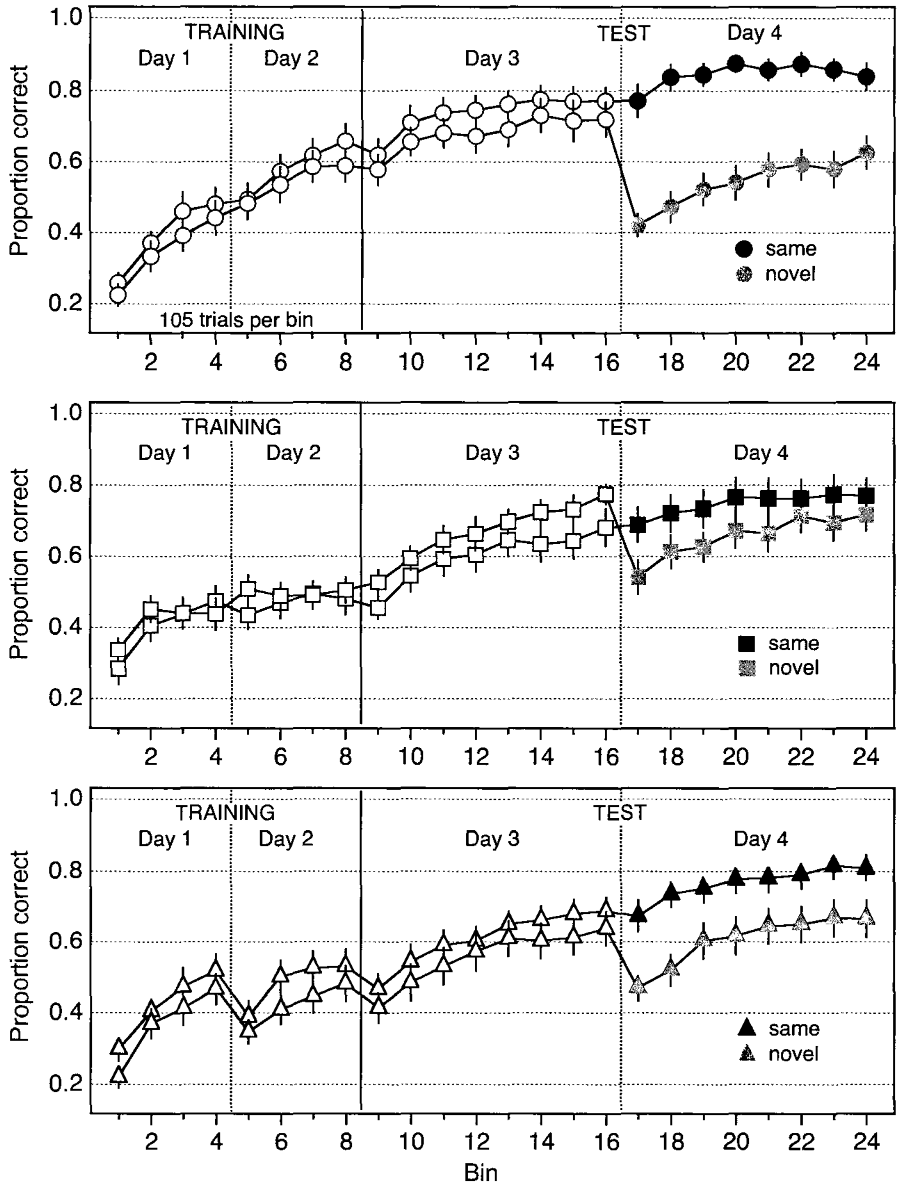


Figure 7.1: Effects on accuracy of three different types of training (Days 1 and 2). On Days 3 & 4 (test) all groups performed the task with a fixed set of textures; half the subjects performed the task with the same fixed set of textures on both days (black), and the other half transfered from a fixed set to a novel set from Day 3 to Day 4 (grey). Top panel: Subjects practiced with the same fixed set of 10 textures on Days 1, 2, & 3. Middle panel: Subjects practiced with a novel set of 10 textures on every trial on Days 1 & 2, therefore they never saw the same texture twice, and viewed in all, 840 sets of 10 textures during the first two days. Bottom panel: Subjects switched between a fixed set of textures from Day 1 to Day 2, and from Day 2 to Day 3; this group therefore had prior experience in switching between a fixed set of textures, and viewed three sets of 10 textures prior to test

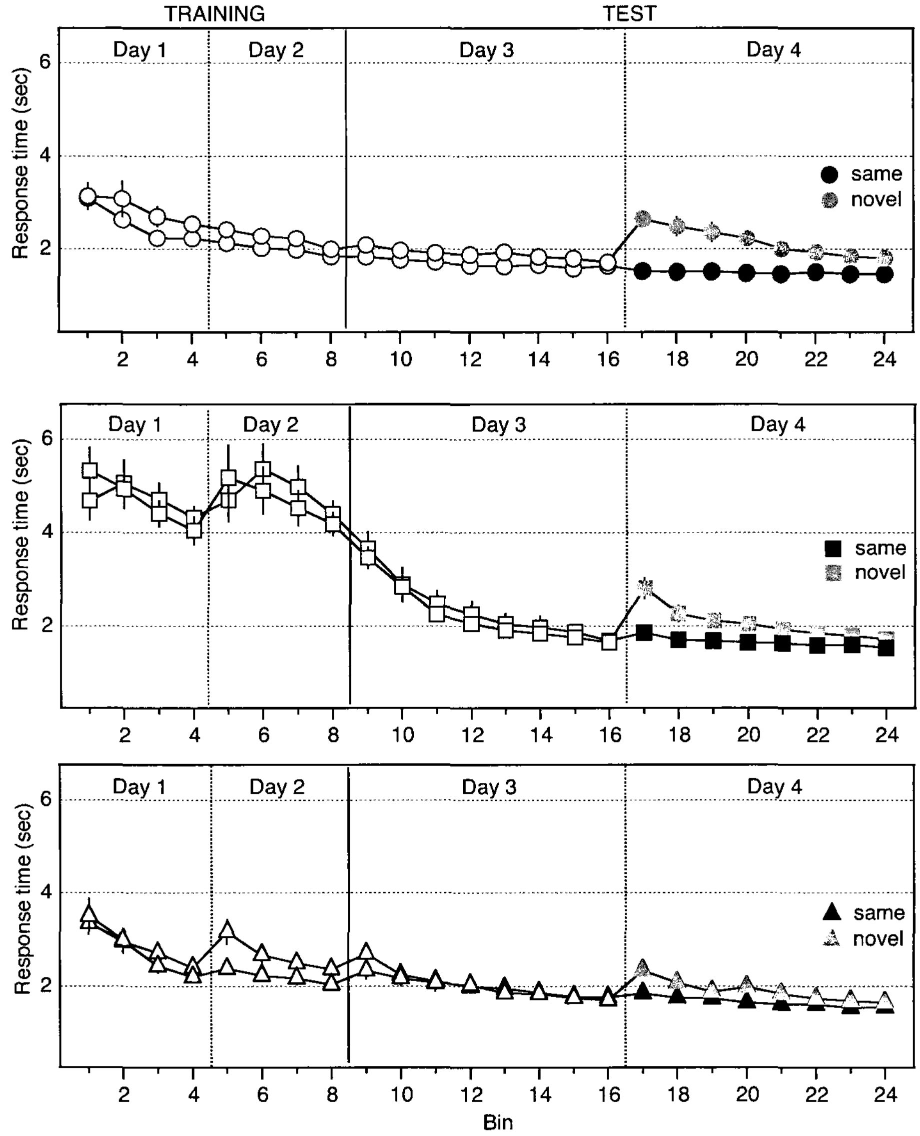


Figure 7.2: Effects on response time of three different types of training (Days 1 and 2). See previous figure for details on the training conditions.

training items. Another way is to increase the sample size of the training items.

Does transfer of learning increase when the sample size of the training set is increased? From the most recent experiment in this series, the answer is yes. When subjects practice 10AFC texture identification in a condition in which a novel set of ten textures is generated on every trial (i.e., when they view a very large pool of textures, and no texture more than once, here called the Unlimited texture condition), they become adept at transferring from one fixed set of textures to another (Figures 7.1 and 7.2, gray trace in middle panels,). Therefore, one way to increase generalization of learning is to increase the sample size of the practice items. The performance of a different group indicates that the previous experience of switching between fixed sets (here called the Switch condition), also facilitates transfer of learning on subsequent occasions (bottom panels, Figures 7.1 & 7.2). Although the Unlimited and Switch conditions produce better generalization of learning, the highest accuracy is achieved by subjects who identify the same set of fixed textures throughout (Fixed condition; black trace in top panels, Figures 7.1 & 7.2). Hence, with high accuracy as a criterion, performance is best after training with a small sample size, but with versatility as a criterion, performance is best after training with a large sample size.

It is interesting that learning occurs at all when no texture is shown more than once. For other tasks, it has been reported that learning in roving conditions - conditions in which several types of stimuli are interleaved during the session - is either abolished (e.g., Adini et al., 2004; Yu et al., 2004; Kuai et al., 2005; Otto et al., 2006), or very gradual, requiring as many as 18,000 trials for any measurable improvements to emerge (Parkosadze et al., 2008). On the other hand, consistent with what we find here, Xiao et al. (2008) recently reported that a double-training paradigm induced complete transfer of perceptual learning across the trained retinal location for the tasks of contrast-discrimination and orientation-discrimination. Paradigm effects on perceptual learning are probably based in perceptual decisions or strategies, but changes to the stimulus representations cannot be entirely ruled out. To address this type of issue, we are now conducting experiments to test whether improved detection of the stimulus affects identification and vice versa.

References

Adini, Y., Wilkonsky, A., Haspel, R., Tsodyks, M., Sagi, D., 2004. Perceptual learning in contrast discrimination: The effect of contrast uncertainty. Journal of Vision 12 (2), 993–1005.

Ahissar, M., Hochstein, S., May 1997. Task difficulty and the specificity of perceptual learning. Nature 387, 401–406.

Beard, B., Levi, D., Reich, L., 1995. Perceptual learning in parafoveal vision. Vision Research 35 (12), 914–924.

Bukach, C. M., Gauthier, I., Tarr, M. J., 2006. Beyond faces and modularity: the power of an expertise framework. Trends Cogn Sci 10 (4), 159–166.

Cox, D. D., DiCarlo, James, J., October 2008. Does learned shape selectivity in inferior temporal cortex automatically generalize across retinal position? The Journal of Neuroscience 28 (40), 10045–10055.

Desimone, R., November 1996. Neural mechanisms for visual memory and their role in attention. Proceedings of the National Academy of Sciences, USA 93, 13494–13499.

Desimone, R., Albright, T., Gross, C., Bruce, C., 1984. Stimulus-selective properties of inferior temporal neurons in the macaque. The Journal of Neuroscience 4 (8), 2051–2062.

Fahle, M., 2005. Perceptual learning: specificity versus generalization. Current Opinion in Neurobiology 15, 154–160.

Fiorentini, A., Berardi, N., 1981. Learning in grating waveform discrimination: specificity for orientation and spatial frequency. Vision Research 21 (7), 1149–1158.

Gauthier, I., Williams, P., Tarr, M. J., Tanaka, J., 1998. Training 'greeble' experts: a framework for studying expert object recognition processes. Vision Res 38 (15-16), 2401–2428.

Gilbert, C., 1994. Early perceptual learning. Proceedings of the National Academy of Sciences of the Unites States of Americe 91, 1195–1197.

Goldstein, A., Chance, J., 1985. Effects of training on japanese face recognition: Reduction of the other-race effect. B Psychonomic Soc 23, 211–214.

Ito, M., Fujita, I., Tamura, H., Tanaka, K., 1994. Processing on contrast polarity of visual images in inferotemporal cortex of macaque monkey. Cerebral Cortex 5, 499–508.

Jagadeesh, B., Chelazzi, L., Mishkin, M., Desimone, R., 2001. Learning increases stimulus salience in anterior inferior temporal cortex of the macaque. Journal of Neurophysiology 86, 290–303.

Karni, A., Bertini, G., 1997. Learning perceptual skills: behavioral probes into adult cortical plasticity. Curr Opin Neurobiol 7 (4), 530–535.

Karni, A., Sagi, D., 1991. Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. Proc Natl Acad Sci U S A 88 (11), 4966–4970.

Kuai, S.-G., Zhang, J.-Y., Klein, S., Levi, D., Yu, C., 2005. The essential role of stimulus temporal patterning in enabling perceptual learning. Nature Neuroscience 8 (11).

Law, C.-T., Gold, J., 2008. Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. Nature Neuroscience 11 (4).

Li, N., DiCarlo, J., September 2008. Unsupervised natural experience rapidly alters invariant object representation in visual cortex. Science 321, 1502–1506.

Liu, Z., Weinshall, D., January 2000. Mechanisms of generalization in perceptual learning. Vision Research 40 (1), 97–109.

Logothetis, N., Pauls, J., Poggio, T., 1995. Shape representation in the inferior temporal cortex of monkeys. Current Biology 5 (5), 552–563.

Matthews, N., Liu, Z., Geesaman, B., Qian, N., 1999. Perceptual learning on orientation and direction discrimination. Vision Research 30 (22), 3692–3701.

McLaren, I., Mackintosh, N., 2000. An elemental model of associative learning: I. latent inhibition and perceptual learning. Animal Learning and Behavior 28 (3), 211–246.

McLaren, I., Mackintosh, N., 2002. Associative learning and elemental representation: Ii. generalization and discrimination. Animal Learning and Behavior 30 (3), 177–200.

Mednick, S. C., Arman, A. C., Boynton, G. M., 2005. The time course and specificity of perceptual deterioration. Proc Natl Acad Sci U S A 102 (10), 3881–3885.

Mollon, J. D., Danilova, M. V., 1996. Three remarks on perceptual learning. Spat Vis 10 (1), 51–58.

Otto, T., Herzog, M., Fahle, M., Zhaoping, L., 2006. Perceptual learning with spatial uncertainties. Vision Research 46, 3223–3233.

Parkosadze, K., Otto, T., Malania, M., Kezeli, A., Herzog, M., 2008. Perceptual learning of bisection stimuli under roving: Slow and largely specific. Journal of Vision 8 (1), 1–8.

Poggio, T., Fahle, M., Edelman, S., 1992. Fast perceptual learning in visual hyperacuity. Science 256 (5059), 1018–21.

Saksida, L., 1999. Effects of similarity and experience on discrimination learning: A nonassociative connectionist model of perceptual learning. Journal of Experimental Psychology: Animal Behavior Processes 25 (3), 308–323.

Schoups, A. A., Vogels, R., Orban, G. A., 1995. Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. J Physiol 483 (Pt 3), 797–810.

Sigman, M., Gilbert, C. D., 2000. Learning to find a shape. Nat Neurosci 3 (3), 264–269.

Song, J., Skoe, E., Wong, P., Kraus, N., October 2008. Plasticity in the adult human auditory brainstem following short-term linguistic training. Journal of Cognitive Neuroscience 20 (10), 1892–1902.

Tanaka, K., Saito, H., Fukada, Y., Moriya, M., 1991. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. Journal of Neurophysiology 66 (1), 170–189.

Valentine, T., Bruce, V., 1986. The effect of race, inversion and encoding activity upon face recognition. Acta Psychol (Amst) 61 (3), 259–273.

Xiao, L., Zhang, J., Wang, R., Klein, S., Levi, D., Yu, C., 2008. Complete transfer of perceptual learning across retinal locations enabled by double training. Current Biology 18 (24), 1922–1926.

Yi, D.-J., Olson, I. R., Chun, M. M., 2006. Shape-specific perceptual learning in a figure-ground segregation task. Vision Res 46 (6-7), 914–924.

Yu, C., Klein, S., Levi, D., March 2004. Perceptual learning in contrast discrimination and the (minimal) role of context. Journal of Vision 4, 169–182.