ON THE ROLE OF HORIZONTAL STRUCTURE IN HUMAN FACE IDENTIFICATION

ON THE ROLE OF HORIZONTAL STRUCTURE IN HUMAN FACE IDENTIFICATION

BY

MATTHEW V. PACHAI, B.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF

PSYCHOLOGY, NEUROSCIENCE & BEHAVIOUR

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Doctor of Philosophy (2015) McMaster University

(Psychology, Neuroscience & Behaviour) Hamilton, Ontario, Canada

TITLE:	ON THE ROLE OF HORIZONTAL STRUCTURE IN HUMAN FACE IDENTIFICATION
AUTHOR:	Matthew V. Pachai B.Sc., (Psychology) McMaster University, Hamilton, Canada
SUPERVISOR:	Dr. Allison Sekuler & Dr. Patrick Bennett
NUMBER OF PAGES:	xxi,161

Abstract

The human visual system must quickly and accurately deploy task-and-object-specific processing to successfully navigate the environment, which suggests several interesting research questions: What is the nature of these strategies? Are they flexible? To what extent is this behaviour optimal given the natural statistics of the environment? In this thesis, I explored these questions using human faces, a complex and dynamic source of socially relevant information that we encounter throughout our lives. Specifically, I conducted several experiments examining the role of horizontally-oriented spatial frequency components in face identification. In Chapter 2, I use computational modelling to demonstrate that the structure conveyed by these components is maximally diagnostic for face identity, and show that selective processing of this structure predicts both face identification performance and the face inversion effect. In Chapter 3, I quantify the bandwidth utilized by human observers and relate this sampling strategy to the information structure of face stimuli. In Chapter 4, I show that the selective sampling described in Chapters 2 and 3 is driven by information from the eyes. Finally, in Chapter 5, I show that the impaired horizontal selectivity associated with face inversion is enhanced by practice identifying inverted faces. Together, these experiments characterize a stimulus with differentially diagnostic information sources that, through experience, becomes selectively processed in a manner

associated with task performance. These results contribute to our understanding of expert object processing and may have implications for observers experiencing face perception deficits.

Acknowledgements

Thank you first and foremost to my supervisors, Allison Sekuler and Pat Bennett, for teaching me to think critically, to consider every possibility, and to follow the data wherever they lead. Their guidance has shaped me into the scientist I am today, and for that I am extremely grateful. Thanks also to Dan Goldreich for his insightful feedback and stimulating discussions during committee meetings, and for teaching me to think about the world probabilistically. Thanks finally to David Shore and Bruce Milliken, who welcomed me into their labs and taught me that cognitive psychology is more than a series of uninteresting box models.

This thesis would not have been possible without Donna Waxman, whose unconditional support has helped me more than I could possibly describe. I'm also grateful to my lab mates for the many stimulating discussions and good times, to all my friends in PNB, who I won't attempt to name for fear of leaving someone out, and to my family for their love and support. Finally, I couldn't have done this without Courtney, who supported me more than anyone en route to this lofty goal. Financial support for my PhD was provided by grants from Natural Sciences and Engineering Research Council (CGS-M, PGS-D) and an Ontario Graduate Scholarship (OGS).

Preface

Chapter 1: General Introduction

Author: Matthew V. Pachai

Chapter 2: Sensitivity to information conveyed by horizontal contours is correlated with face identification accuracy

Authors: Matthew V. Pachai, Allison B. Sekuler, & Patrick J. Bennett

Contributions: MVP, ABS, and PJB conceived and designed the study. MVP and PJB conducted the analyses. MVP and PJB conducted the ideal observer analyses. MVP wrote the manuscript with revisions from ABS and PJB. Reprinted from the manuscript published by *Frontiers in Psychology*.

Chapter 3: Effects of bandwidth on selectivity for horizontal structure during face identification

Authors: Matthew V. Pachai, Allison B. Sekuler, & Patrick J. Bennett

Contributions: MVP conceived and designed the study. MVP conducted the analysis with suggestions from ABS and PJB. MVP and PJB conducted the ideal observer analyses. MVP wrote the manuscript with revisions from ABS and PJB.

Chapter 4: Masking of individual facial features reveals the use of horizontal structure in the eyes

Authors: Matthew V. Pachai, Allison B. Sekuler, & Patrick J. Bennett

Contributions: MVP conceived and designed the study. MVP conducted the analysis. MVP conducted the ideal observer analyses. MVP wrote the manuscript with revisions from ABS and PJB.

Chapter 5: The effect of training with inverted faces on the selective use of horizontal structure

Authors: Matthew V. Pachai, Allison B. Sekuler, & Patrick J. Bennett

Contributions: MVP conceived and designed the study. MVP conducted the analysis with suggestions from ABS and PJB. MVP conducted the ideal observer analyses with assistance from PJB. MVP wrote the manuscript with revisions from ABS and PJB.

Chapter 6: General Discussion

Author: Matthew V. Pachai

Contents

List of Figures

Chapter 1

General Introduction

1.1 Face Identification

Every day, human observers encounter a multitude of faces. Although these faces share a similar structure, with two vertically-aligned eyes positioned above a horizontally-aligned nose and mouth, subtle variations between exemplars are sufficient to permit accurate discrimination of identity under highly discrepant size, viewpoint, and lighting conditions. That identification is accurate under such difficult conditions has led many to suggest that face processing is qualitatively unique. In other words, through an innate mechanism (e.g., Farah *et al.*, 1998; Kanwisher, 2000) or a lifetime of experience (e.g., Gauthier *et al.*, 1998), face processing is characterized by mechanisms that are not deployed to process other categories of objects (Rossion, 2008). Alternatively, face processing could be qualitatively similar to the processing of other objects, characterized only by its specialization (Riesenhuber *et al.*, 2004; Sekuler *et al.*, 2004). In either event, understanding the nature of face processing can provide great insight into the general principles of expert object processing in the visual system. In this

introduction, I will describe previous research directed toward this goal, which was the motivation for each the experiments undertaken in this thesis.

1.2 Face Inversion Effects

To explore the hypothesis that face perception is qualitatively unique, one might compare identification of faces to that of other objects. However, there are several potential confounds in such a design. Most importantly, faces may be more discriminable than other object categories, confounding processing differences with task difficulty. This problem could be compounded by an inappropriate non-face control: each possible control category (e.g., cars, houses, animals) could be more or less discriminable than the others, or itself engage specialized processing that confounds a comparison with faces. One control category that does not suffer from these confounds is the inverted face; upright and inverted faces are equally discriminable, and inverted faces are unlikely to engage specialized processing because of their rarity in everyday environments. For these reasons, inverted faces have become the most popular control stimulus in studies of face expertise.

It has long been known that inverted faces are difficult to identify (Hochberg and Galper, 1967). However, it was Yin (1969) who first demonstrated that picture-plane inversion disproportionately impairs the perception of faces, relative to other objects. In his experiments, an upright or inverted face or object was briefly presented, followed by two exemplars of that category. When asked to identify the previously-presented item, participants were significantly more impaired by inversion of faces than other objects. Following Yin's seminal work, the face inversion effect has been replicated robustly across a variety of task and presentation conditions (e.g., Barton *et al.*,

2001; Eimer, 2000; Gaspar *et al.*, 2008b; Goffaux and Rossion, 2007; Hills *et al.*, 2011; Hussain *et al.*, 2009; Itier and Taylor, 2002; LeGrand *et al.*, 2001; Mondloch *et al.*, 2002; Persike *et al.*, 2014; Riesenhuber *et al.*, 2004; Russell and Sinha, 2007; Sekuler *et al.*, 2004; Yovel and Kanwisher, 2004). In addition to being less accurate, inverted face processing is now known to be slower (e.g., Jacques and Rossion, 2007; Jacques *et al.*, 2007; Meinhardt-Injac *et al.*, 2010; Rossion *et al.*, 2000) and associated with inappropriate sampling strategies (e.g., Barton *et al.*, 2006; de Lissa *et al.*, 2014; Hills *et al.*, 2011 but see Williams and Henderson, 2007). Together, these results demonstrate that upright faces are processed differently than inverted faces, but do not clarify whether these differences reflect a qualitatively different mechanism.

1.3 Holistic Processing

In his original manuscript, Yin (1969) argued that face inversion effects are driven by two performance impairments: one that affects all objects, and one that affects faces only. More specifically, Yin (1969) suggested that upright faces engage a holistic processing strategy that is not deployed for inverted faces or other objects. Since Yin's seminal work, this notion has become so accepted that inversion effects have been taken as evidence of holistic processing in other domains (e.g., Bernard *et al.*, 2012; Reed *et al.*, 2003). However, it is not clear that inversion effects necessarily demonstrate holistic processing (Valentine, 1988), nor is it clear how holistic processing should be operationalized.

Young *et al.* (1987) were among the first to operationalize holistic processing in an experimental framework. In their experiments, Young *et al.* combined the top half of one face with the bottom half of another, presenting these composite stimuli

either aligned or misaligned (i.e., offset horizontally). When the two face-halves were misaligned, observers were faster and more accurate to identify the top halves than when the face-halves were aligned. This so-called Composite Face Effect (CFE) was taken as evidence that observers obligatorily deploy holistic processes to identify intact faces, as they could not disregard the mismatched bottom halves. Critically, Young *et al.* (1987) also found that inverted faces produced a significantly smaller CFE, leading them to suggest that holistic processing is disrupted by face inversion. Following this demonstration, the CFE has been used to explore whether holistic processing is present in children (de Heering *et al.*, 2007; Mondloch *et al.*, 2007; Turati *et al.*, 2010), older adults (Boutet and Faubert, 2006; Konar *et al.*, 2013), or patients with prosopagnosia (Avidan *et al.*, 2011; Busigny and Rossion, 2010; Palermo *et al.*, 2011), and to understand the nature of face-specific neural responses (Jacques and Rossion, 2009; Schiltz *et al.*, 2010). Indeed, the composite effect is so accepted that it has been used to explore holistic processing of non-face stimuli (e.g. Soria Bauser *et al.*, 2011, 2014; Willems *et al.*, 2014).

Studies using the CFE as a tool to investigate face processing make two critical assumptions: that the CFE is a valid measure of holistic processing, and that holistic processing supports face identification accuracy. If both of these assumptions are appropriate, then the magnitude of the CFE should be correlated with face identification accuracy. In the first study to explore this hypothesis, Konar *et al.* (2010) found no correlation between the CFE, as measured either by accuracy or reaction time, and face identification accuracy. Indeed, subsequent experiments have robustly observed this correlation to be small or non-significant (Konar *et al.*, 2013; Richler *et al.*, 2011; Wang *et al.*, 2012), and contingent on particular experimental details

(e.g., Boutet and Faubert, 2006; Cheung *et al.*, 2008; DeGutis *et al.*, 2013; Richler *et al.*, 2011). These studies cast doubt on the validity of the CFE as a measure of holistic processing, and indeed on the notion that differences in holistic processing underlie the face inversion effect. For these reasons, a different approach may be required to explore the nature of upright face identification expertise.

1.4 Selective Processing

1.4.1 Spatial Domain

Human face perception is well characterized by selective spatial processing. For example, Yarbus (1967) famously showed that observers exhibit a stereotypically T-shaped pattern of fixations when viewing a face, allocating attention to each of the eyes, nose, and mouth in turn. We know these fixations facilitate encoding of identity, as restricting observers to a central fixation during face learning severely hinders subsequent recognition (Henderson *et al.*, 2005) and fixation patterns are atypical in disorders associated with identification deficits (Bukach *et al.*, 2006, 2008; Hernandez *et al.*, 2009; Klin *et al.*, 2002; Lê *et al.*, 2003; Pelphrey *et al.*, 2002; Schwarzer *et al.*, 2007; Xivry *et al.*, 2008). However, fixation patterns may not necessarily reveal the spatial characteristics of information sampling during face identification, because information collected during a fixation may not be used to support an identification decision. For example, Hsiao and Cottrell (2008) demonstrated that only one or two fixations are sufficient to support old-new face discrimination. In their experiment, Hsiao and Cottrell presented faces in the visual periphery and permitted observers to make one, two, three, or four fixations to any location prior to their decision. They found that

permitting more than two fixations conferred no further benefit to performance. Using a face identification paradigm in which observers were only permitted one saccade, Peterson and Eckstein (2012) demonstrated that observers consistently fixated just above the nose, between the two eyes. With computational modelling techniques, Peterson and Eckstein revealed that this location is optimal for a single fixation, because it permits collection of information from both eyes. That this information is collected from outside foveal vision is further evidence that saccade targets may not quantify information sampling during face identification tasks.

Although eye tracking techniques may have several limitations, psychophysical techniques can supplement these results. In the classification image technique (Ahumada Jr, 1996), each pixel of a visual stimulus is degraded by white noise. If we assume that the noise randomly alters the stimulus to look more or less like a particular alternative on every trial, the classification image technique quantifies the information supporting visual decisions by correlating the noise in each pixel with an observer's responses. If the correlation at a particular location is non-zero, we have evidence that the observer was using information from that location in their decision. Using similar logic, the bubbles technique (Gosselin and Schyns, 2001) involves filtering the visual stimulus with randomly distributed spatial apertures that make available a subset of the total information on a given trial. By correlating observer responses with information availability at each spatial location, the bubbles technique quantifies spatial sampling during a perceptual task. Studies using classification images (Mangini and Biederman, 2004; Sekuler *et al.*, 2004) and bubbles (Gosselin and Schyns, 2001; Vinette *et al.*, 2004) reveal that observers use exclusively the eyes and eyebrows to identify faces. Further, Gold *et al.* (2004) showed that

observers use the eyes more exclusively following perceptual learning, and Caldara *et al.* (2005) showed that observers with prosopagnosia use the eyes less exclusively than controls. Together, these results demonstrate that information from the eyes underlies face identification performance. However, selective processing of information from the eyes may not uniquely characterize upright face identification expertise. Recent studies have shown that observers also sample from the eyes of inverted faces, despite having no expertise with these stimuli (Sekuler *et al.*, 2004; Williams and Henderson, 2007). Indeed, Gaspar *et al.* (2008b) showed that observers simply process information from inverted faces less efficiently than upright, but how to quantify the information available in faces remains unclear, as does the specific effect of inversion on sampling this information. However, a different approach may yield answers to these questions.

1.4.2 Fourier Domain

It has long been known that visual images are represented by activity in V1 cells tuned to orientation and spatial frequency information conveyed at specific spatial locations (De Valois *et al.*, 1982; DeValois and DeValois, 1988; Hubel and Wiesel, 1968). Together, these cells comprise pathways, or channels, that approximate the output of a two-dimensional Fourier decomposition. Many classic models of perception suggest that the visual system treats each of these pathways as independent sources of information, combining their output in a manner optimal for the task at hand (e.g., Campbell and Robson, 1968; Sachs *et al.*, 1971; Thomas *et al.*, 1993). Although these tuned pathways may not always be independent (e.g., Olzak and Thomas, 1991, 1992), such models have provided a powerful framework to explore

information processing during visual discrimination tasks such as face identification.

Considerable evidence suggests that object recognition processes are supported by narrow-band spatial frequency channels (e.g. Braje *et al.*, 1995; Fiorentini and Berardi, 1981; Fiorentini *et al.*, 1983; Solomon and Pelli, 1994; Tjan *et al.*, 1995), and face identification in particular seems to rely on a band of spatial frequencies around 8-10 cycles per face width (Boutet *et al.*, 2003; Costen *et al.*, 1996; Gold *et al.*, 1999; Hayes *et al.*, 1986; Näsänen, 1999; Peli *et al.*, 1994). We also know that faces contain information at all spatial scales (Gold *et al.*, 1999), and experience can alter spatial frequency selectivity (e.g., Fiorentini and Berardi, 1980; Hua *et al.*, 2010). For these reasons, one may hypothesize that spatial frequency selectivity reflects experience with upright faces, likely gained during early childhood (Gao and Maurer, 2011). However, recent studies have demonstrated nearly identical spatial frequency selectivity for upright and inverted faces (Gaspar *et al.*, 2008a; Willenbockel *et al.*, 2010), suggesting that spatial frequency selectivity does not uniquely characterize upright face identification expertise. However, object recognition processes are also supported by orientation-selective channels (e.g., Thomas and Gille, 1979; Taylor *et al.*, 2014), and in this thesis I will explore the extent to which selective sampling of information conveyed by orientation channels characterizes human face identification.

1.5 Thesis Experiments

Dakin and Watt (2009) first demonstrated that information conveyed by horizontally-oriented spatial frequency components is a diagnostic cue to face identity. Specifically, Dakin and Watt measured identification accuracy for band-pass filtered famous faces retaining a subset of all orientation components. They observed higher accuracy

when horizontal components were retained (approx. 55%) than when vertical components were retained (approx. 35%). They further demonstrated using computational analyses that faces, relative to other objects, contain systematically clustered horizontal structure that uniquely conveys face identity. This structure, they argued, is tolerant to manipulations that do not typically affect identification (e.g., pose variation, distortion, luminance noise) and disrupted by manipulations known to affect identification (e.g., contrast polarity reversal, picture-plane inversion). This research inspired the series of experiments comprising the present thesis.

In my first experiment (Chapter 2), I explored the relationship between selective processing of horizontal orientation components and face identification. I showed that reliance on horizontal information, relative to vertical, is significantly correlated with face identification performance in a baseline condition, that selective reliance on horizontal information is abolished by face inversion, and that the extent of this abolition is correlated with the face inversion effect. In my next experiment (Chapter 3), I quantified more precisely the diagnostic orientation information available in face stimuli using an ideal observer analysis (i.e., a mathematical algorithm that utilizes optimally the available stimulus information). I showed that diagnostic horizontal structure is conveyed by a relatively narrow channel, but that human observers utilize a much larger bandwidth to achieve accurate identification performance. In Chapter 4, I present the results of an experiment examining the link between the spatial selectivity and horizontal selectivity observed during upright face identification, demonstrating that human face identification is supported by horizontal structure conveyed near the eyes. In my final experiment (Chapter 5), I showed that perceptual training with inverted faces ameliorates the face inversion effect by improving

selective processing of horizontal structure from inverted faces. Finally, in Chapter 6, I provide a general discussion that ties together these results to explore more generally how horizontal structure supports human face identification and suggests several future directions for this line of research. Collectively, these chapters explore a source of information in face stimuli whose selective processing characterizes the nature of human face identification expertise.

Bibliography

Ahumada Jr, A. J. (1996). Perceptual classification images from Vernier acuity masked by noise. *Perception ECVP abstract*, **25**, 0.

Avidan, G., Tanzer, M., and Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. *Neuropsychologia*, **49**(9), 2541–2552.

Barton, J. J. S., Cherkasova, M., and O'Connor, M. (2001). Covert recognition in acquired and developmental prosopagnosia. *Neurology*, **57**(7), 1161–8.

Barton, J. J. S., Radcliffe, N., Cherkasova, M. V., Edelman, J., and Intriligator, J. M. (2006). Information processing during face recognition: The effects of familiarity, inversion, and morphing on scanning fixations. *Perception*, **35**(8), 1089–1105.

Bernard, P., Gervais, S. J., Allen, J., Campomizzi, S., and Klein, O. (2012). Integrating sexual objectification with object versus person recognition: the sexualized-body-inversion hypothesis. *Psychological science*, **23**(5), 469–71.

Boutet, I. and Faubert, J. (2006). Recognition of faces and complex objects in younger and older adults. *Memory & cognition*, **34**(4), 854–64.

Boutet, I., Collin, C. A., and Faubert, J. (2003). Configural face encoding and spatial frequency information. *Perception & Psychophysics*, **65**(7), 1078–1093.

Braje, W. L., Tjan, B. S., and Legge, G. E. (1995). Human efficiency for recognizing and detecting low-pass filtered objects. *Vision Research*, **35**(21), 2955–66.

Bukach, C. M., Bub, D. N., Gauthier, I., and Tarr, M. J. (2006). Perceptual expertise effects are not all or none: spatially limited perceptual expertise for faces in a case of prosopagnosia. *Journal of cognitive neuroscience*, **18**(1), 48–63.

Bukach, C. M., Le Grand, R., Kaiser, M. D., Bub, D. N., and Tanaka, J. W. (2008). Preservation of mouth region processing in two cases of prosopagnosia. *Journal of Neuropsychology*, **2**(1), 227–244.

Busigny, T. and Rossion, B. (2010). Acquired prosopagnosia abolishes the face inversion effect. *Cortex*, **46**(8), 965–81.

Caldara, R., Schyns, P. G., Mayer, E., Smith, M. L., Gosselin, F., and Rossion, B. (2005). Does prosopagnosia take the eyes out of face representations? Evidence for a defect in representing diagnostic facial information following brain damage. *Journal of cognitive neuroscience*, **17**(10), 1652–66.

Campbell, F. W. and Robson, J. G. (1968). Application of fourier analysis to the visibility of gratings. *Journal of Physiology*, **197**, 551–566.

Cheung, O. S., Richler, J. J., Palmeri, T. J., and Gauthier, I. (2008). Revisiting the role of spatial frequencies in the holistic processing of faces. *Journal of experimental psychology. Human perception and performance*, **34**(6), 1327–1336.

Costen, N. P., Parker, D. M., and Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception & psychophysics*, **58**(4), 602–612.

Dakin, S. C. and Watt, R. J. (2009). Biological bar codes in human faces. *Journal of Vision*, **9**(4), 1–10.

de Heering, A., Houthuys, S., and Rossion, B. (2007). Holistic face processing is mature at 4 years of age: evidence from the composite face effect. *Journal of experimental child psychology*, **96**(1), 57–70.

de Lissa, P., McArthur, G., Hawelka, S., Palermo, R., Mahajan, Y., and Hutzler, F. (2014). Fixation location on upright and inverted faces modulates the N170. *Neuropsychologia*, **57**, 1–11.

De Valois, R. L., Albrecht, D. G., and Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, **22**(5), 545–59.

DeGutis, J., Wilmer, J., Mercado, R. J., and Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition*, **126**(1), 87–100.

DeValois, R. L. and DeValois, K. K. (1988). *Spatial vision*. Number 14. Oxford University Press, New York.

Eimer, M. (2000). Effects of face inversion on the structural encoding and recognition of faces. Evidence from event-related brain potentials. *Cognitive brain research*, **10**(1-2), 145–58.

Farah, M. J., Wilson, K. D., Drain, H. M., and Tanaka, J. W. (1998). What is "special" about face perception? *Psychological review*, **105**(3), 482–98.

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

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

Fiorentini, A., Maffei, L., and Sandini, G. (1983). The role of high spatial frequencies in face perception. *Perception*, **12**(2), 195–201.

Gao, X. and Maurer, D. (2011). A comparison of spatial frequency tuning for the recognition of facial identity and facial expressions in adults and children. *Vision research*, **51**(5), 508–19.

Gaspar, C. M., Sekuler, A. B., and Bennett, P. J. (2008a). Spatial frequency tuning of upright and inverted face identification. *Vision Research*, **48**(28), 2817–26.

Gaspar, C. M., Bennett, P. J., and Sekuler, A. B. (2008b). The effects of face inversion and contrast-reversal on efficiency and internal noise. *Vision Research*, **48**(8), 1084–95.

Gauthier, I., Williams, P., Tarr, M. J., and Tanaka, J. W. (1998). Training 'greeble' experts: a framework for studying expert object recognition processes. *Vision research*, **38**(15-16), 2401–28.

Goffaux, V. and Rossion, B. (2007). Face inversion disproportionately impairs the perception of vertical but not horizontal relations between features. *Journal of Experimental Psychology: Human Perception and Performance*, **33**(4), 995–1002.

Gold, J. M., Bennett, P. J., and Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research*, **39**(21), 3537–60.

Gold, J. M., Sekuler, A. B., and Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science*, **28**(2), 167–207.

Gosselin, F. and Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision research*, **41**(17), 2261–71.

Hayes, T., Morrone, M. C., and Burr, D. C. (1986). Recognition of positive and negative bandpass-filtered images. *Perception*, **15**(5), 595–602.

Henderson, J. M., Williams, C. C., and Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & cognition*, **33**(1), 98–106.

Hernandez, N., Metzger, A., Magné, R., Bonnet-Brilhault, F., Roux, S., Barthelemy, C., and Martineau, J. (2009). Exploration of core features of a human face by healthy and autistic adults analyzed by visual scanning. *Neuropsychologia*, **47**(4), 1004–1012.

Hills, P. J., Ross, D. A., and Lewis, M. B. (2011). Attention misplaced: the role of diagnostic features in the face-inversion effect. *Journal of experimental psychology. Human perception and performance*, **37**(5), 1396–406.

Hochberg, J. E. and Galper, R. E. (1967). Recognition of Faces: I. An Exploratory Study. *Psychonomic Science*, **9**(12), 619–620.

Hsiao, J. H.-w. and Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological science*, **19**(10), 998–1006.

Hua, T., Bao, P., Huang, C.-B., Wang, Z., Xu, J., Zhou, Y., and Lu, Z.-L. (2010). Perceptual Learning Improves Contrast Sensitivity of V1 Neurons in Cats. *Current biology*, **20**(10), 887–894.

Hubel, D. H. and Wiesel, T. N. (1968). Receptive Fields and Functional Architecture of Monkey Striate Cortex. *Journal of Physiology*, **195**, 215–243.

Hussain, Z., Sekuler, A. B., and Bennett, P. J. (2009). Perceptual learning modifies inversion effects for faces and textures. *Vision research*, **49**(18), 2273–84.

Itier, R. J. and Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: a repetition study using ERPs. *NeuroImage*, **15**(2), 353–72.

Jacques, C. and Rossion, B. (2007). Early electrophysiological responses to multiple face orientations correlate with individual discrimination performance in humans. *NeuroImage*, **36**(3), 863–76.

Jacques, C. and Rossion, B. (2009). The initial representation of individual faces in the right occipito-temporal cortex is holistic: electrophysiological evidence from the composite face illusion. *Journal of vision*, **9**(6), 8.1–16.

Jacques, C., D'Arripe, O., and Rossion, B. (2007). The time course of the inversion effect during individual face discrimination. *Journal of vision*, **7**(8), 3.

Kanwisher, N. G. (2000). Domain specificity in face perception. *Nature Neuroscience*, **3**(8), 759–763.

Klin, A., Jones, W., Schultz, R., Volkmar, F., and Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of general psychiatry*, **59**(9), 809–16.

Konar, Y., Bennett, P. J., and Sekuler, A. B. (2010). Holistic Processing Is Not

Correlated With Face-Identification Accuracy. *Psychological Science*, **21**(1), 38–43.

Konar, Y., Bennett, P. J., and Sekuler, A. B. (2013). Effects of aging on face identification and holistic face processing. *Vision research*, **88**, 38–46.

Lê, S., Raufaste, E., and Démonet, J. F. (2003). Processing of normal, inverted, and scrambled faces in a patient with prosopagnosia: Behavioural and eye tracking data. *Cognitive Brain Research*, **17**(1), 26–35.

LeGrand, R., Mondloch, C. J., Maurer, D., and Brent, H. P. (2001). Early visual experience and face processing. *Science*, **410**, 890.

Mangini, M. and Biederman, I. (2004). Making the ineffable explicit: estimating the information employed for face classifications. *Cognitive Science*, **28**(2), 209–226.

Meinhardt-Injac, B., Persike, M., and Meinhardt, G. (2010). The time course of face matching by internal and external features: Effects of context and inversion. *Vision Research*, **50**(16), 1598–1611.

Mondloch, C. J., Le Grand, R., and Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, **31**(5), 553–566.

Mondloch, C. J., Pathman, T., Maurer, D., Le Grand, R., and de Schonen, S. (2007). The composite face effect in six-year-old children: Evidence of adult-like holistic face processing. *Visual Cognition*, **15**(5), 564–577.

Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research*, **39**(23), 3824–3833.

Olzak, L. A. and Thomas, J. P. (1991). When Orthogonal Processed Orientations Independently. *Vision Research*, **31**(1), 51–57.

Olzak, L. A. and Thomas, J. P. (1992). Configural effects constrain Fourier models of pattern discrimination. *Vision research*, **32**(10), 1885–1898.

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., and Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. *Neuropsychologia*, **49**(5), 1226–1235.

Peli, E., Lee, E., Trempe, C. L., and Buzney, S. (1994). Image enhancement for the visually impaired: the effects of enhancement on face recognition. *Journal of the Optical Society of America. A, Optics, image science, and vision*, **11**(7), 1929–1939.

Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., and Piven, J. (2002). Visual Scanning of Faces in Autism. *Journal of Autism and Developmental Disorders*, **32**(4), 249–261.

Persike, M., Meinhardt-Injac, B., and Meinhardt, G. (2014). The face inversion effect in opponent-stimulus rivalry. *Frontiers in Human Neuroscience*, **8**(May), 1–11.

Peterson, M. F. and Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences*, **109**(48).

Reed, C. L., Stone, V. E., Bozova, S., and Tanaka, J. (2003). The body-inversion effect. *Psychological Science*, **14**(4), 302–308.

Richler, J. J., Cheung, O. S., and Gauthier, I. (2011). Beliefs alter holistic face processing I if response bias is not taken into account. *Journal of Vision*, **11**, 1–13.

Riesenhuber, M., Jarudi, I., Gilad, S., and Sinha, P. (2004). Face processing in humans is compatible with a simple shape-based model of vision. *Proceedings of the Royal Society B*, **271 Suppl**, S448–50.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta psychologica*, **128**(2), 274–89.

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., and Crommelinck, M. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. *Neuroreport*, **11**(1), 69–74.

Russell, R. and Sinha, P. (2007). Real-world face recognition: The importance of surface reflectance properties. *Perception*, **36**(9), 1368–1374.

Sachs, M. B., Nachmias, J., and Robson, J. G. (1971). Spatial-frequency channels in human vision. *Journal of the Optical Society of America*, **61**(9), 1176–86.

Schiltz, C., Dricot, L., Goebel, R., and Rossion, B. (2010). Holistic perception of individual faces in the right middle fusiform gyrus as evidenced by the composite face illusion. *Journal of vision*, **10**(2), 25.1–16.

Schwarzer, G., Huber, S., Grüter, M., Grüter, T., Gross, C., Hipfel, M., and Kennerknecht, I. (2007). Gaze behaviour in hereditary prosopagnosia. *Psychological research*, **71**(5), 583–90.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., and Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, **14**(5), 391–6.

Solomon, J. A. and Pelli, D. G. (1994). The visual filter mediating letter identification. *Nature*, **369**, 395–397.

Soria Bauser, D. A., Suchan, B., and Daum, I. (2011). Differences between perception of human faces and body shapes: Evidence from the composite illusion. *Vision Research*, **51**(1), 195–202.

Soria Bauser, D. A., Schriewer, E., and Suchan, B. (2014). Dissociation between the behavioural and electrophysiological effects of the face and body composite illusions. *British Journal of Psychology*, **106**, 414–432.

Taylor, C. P., Bennett, P. J., and Sekuler, A. B. (2014). Evidence for adjustable bandwidth orientation channels. *Frontiers in Psychology*, **5**, 1–10.

Thomas, J. P. and Gille, J. (1979). Bandwidths of orientation channels in human vision. *Journal of the Optical Society of America*, **69**(5), 652–660.

Thomas, J. P., Olzak, L. A., and Shimozaki, S. S. (1993). The role of Fourier components in discrimination between two types of plaid patterns. *Vision research*, **33**(11), 1573–1579.

Tjan, B. S., Braje, W. L., Legge, G. E., and Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research*, **35**(21), 3053–69.

Turati, C., Di Giorgio, E., Bardi, L., and Simion, F. (2010). Holistic Face Processing in Newborns, 3-Month-Old Infants, and Adults: Evidence From the Composite Face Effect. *Child Development*, **81**(6), 1894–1905.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, **79**(4), 471–491.

Vinette, C., Gosselin, F., and Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: its in the eyes. *Cognitive Science*, **28**(2), 289–301.

Wang, R., Li, J., Fang, H., Tian, M., and Liu, J. (2012). Individual differences in holistic processing predict face recognition ability. *Psychological science*, **23**(2), 169–77.

Willems, S., Vrancken, L., Germeys, F., and Verfaillie, K. (2014). Holistic processing of human body postures: Evidence from the composite effect. *Frontiers in Psychology*, **5**, 1–9.

Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J. W., Bub, D. N., and Gosselin, F. (2010). Does face inversion change spatial frequency tuning? *Journal of experimental psychology. Human perception and performance*, **36**(1), 122–35.

Williams, C. C. and Henderson, J. M. (2007). The face inversion effect is not a consequence of aberrant eye movements. *Memory & cognition*, **35**(8), 1977–1985.

Xivry, J.-J. O., Ramon, M., Lefèvre, P., and Rossion, B. (2008). Reduced fixation on the upper area of personally familiar faces following acquired prosopagnosia. *Journal of Neuropsychology*, **2**(1), 245–268.

Yarbus, A. L. (1967). *Eye Movements During Perception of Complex Objects*. Plenum Press, New York.

Yin, R. K. (1969). Looking at Upside-Down Faces. *Journal of Experimental Psychology*, **81**(1), 141–145.

Young, A. W., Hellawell, D., and Hay, D. C. (1987). Configurational information in face perception. *Perception*, **16**(6), 747–59.

Yovel, G. and Kanwisher, N. (2004). Face perception: domain specific, not process specific. *Neuron*, **44**(5), 889–98.

Chapter 2

Sensitivity to information conveyed by horizontal contours is correlated with face identification accuracy

2.1 Introduction

We detect, discriminate, and recognize hundreds of faces every day. However, despite the apparent ease with which face recognition normally operates, there are some conditions in which we experience difficulty. For example, rotating a face 180 deg in the picture plane significantly impairs recognition, and these effects of rotation appear to be larger for faces than for other kinds of objects (Yin, 1969; Valentine, 1988; Husk *et al.*, 2007). This well-established face inversion effect is interesting because the physical information available to discriminate two inverted faces is the same as that available to discriminate two upright faces, and therefore a difference in perceptual processing or observer strategies must underlie the face inversion effect.

The cause of the inversion effect remains a matter of debate. One of the most commonly held theories is that upright and inverted faces are processed using qualitatively different mechanisms: with holistic/configural mechanisms dominating for upright, but not inverted faces (Diamond and Carey, 1986; Young *et al.*, 1987; Tanaka and Farah, 1993; Farah *et al.*, 1995; Rossion, 2008). However, it has also been suggested that upright and inverted face processing differs quantitatively, not qualitatively: in effect, that upright faces are processed more efficiently than inverted faces (Valentine, 1988; Sekuler *et al.*, 2004; Riesenhuber *et al.*, 2004; Yovel and Kanwisher, 2004). Using the classification image technique, Sekuler *et al.* (2004) found that observers relied on information carried by pixels near the eyes and eyebrows to identify both upright and inverted faces. Based on this result, Sekuler *et al.* suggested that observers use similar spatial regions to identify upright and inverted faces, but that inversion produces a quantitative difference in the ability to extract relevant information from those regions. Gaspar *et al.* (2008b) tested this hypothesis directly using the equivalent noise paradigm, and found that inversion decreased calculation efficiency alone, supporting the notion that observers simply use available physical information in the stimuli less effectively when processing inverted faces.

What then leads to decreased processing efficiency for inverted faces compared to upright faces? One possibility is that different spatial frequencies are used to identify upright and inverted faces; however, direct comparisons of the spatial frequency tuning of upright and inverted face identification reveal that observers rely on similar spatial frequencies in both cases (Gaspar *et al.*, 2008a; Willenbockel *et al.*, 2010). It is unlikely, therefore, that the face inversion effect is caused by observers using different bands of spatial frequencies to identify upright and inverted faces.

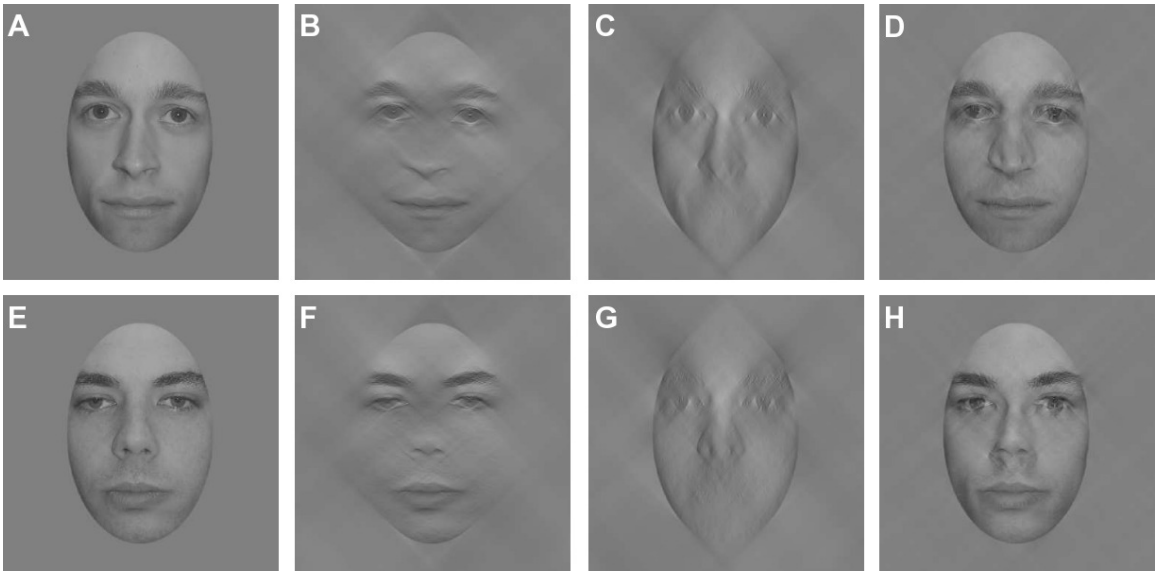


Figure 2.1: Two faces (A,E) filtered to retain only horizontal (B,F) or vertical (C,G) information (half-bandwidth = 45 deg). Hybrid faces (D,H) constructed using horizontal information from one face and vertical information from the other resemble the face from which the horizontal information is drawn (D=B+G and H=F+C). Note that equating the RMS contrast of the filtered components has a negligible effect on the hybrids.

Spatial frequency selectivity is, of course, only one way that identification may differ for upright and inverted faces. Recently, Dakin and Watt (2009) demonstrated that orientation information, specifically conveyed by horizontal contours, may be especially useful for face identification (see also Figure 2.1). Given this finding, perhaps a difference in the use of horizontal information may explain the performance deficits incurred following face inversion. Goffaux and Dakin (2010) examined this hypothesis with face stimuli filtered to contain narrow bands of orientations centred on horizontal, vertical, or both orientations. Using a same/different paradigm, Goffaux and Dakin found that performance was better for upright faces containing horizontal information than upright faces containing vertical information. However, when the faces were inverted, overall performance decreased and the horizontal advantage disappeared. This result demonstrates that discrimination of upright faces is indeed supported by the use of horizontal information, but this orientation difference disappears when the face is inverted. In a series of additional experiments, Goffaux and Dakin (2010) demonstrated the importance of horizontal information for other face phenomena such as identity aftereffects, viewpoint-invariance, and holistic processing. However, it remains unclear whether the the importance of horizontal contours reflects the additional diagnostic information conveyed by that orientation (Dakin and Watt, 2009) or by observers processing that orientation more efficiently.

The primary goal of this study was to disentangle the preferential use of horizontal information by human observers from the informational structure of the stimulus. To this end, we employed a 10-AFC face identification task in which the stimuli were masked with orientation-filtered Gaussian noise. This approach allowed us to assess the importance of different orientation bands by measuring the decrement in

identification performance incurred when they are masked. We also used an ideal observer analysis to systematically assess the information available at each orientation band. An ideal observer is an optimal decision maker that achieves the best possible performance on a task given the available stimulus information (Bennett and Banks, 1987; Banks *et al.*, 1991; Tjan *et al.*, 1995). By measuring the effect of orientation-filtered noise on ideal performance, we obtained measures of how much diagnostic information is carried by different bands of orientations, which in turn allowed us to estimate how efficiently human observers used the available information at different orientations.

Based on the findings of Dakin and Watt (2009), we predicted that the ideal observer would show larger masking effects for horizontally-oriented noise with a decrease in masking at off-horizontal orientations. Moreover, based on the findings of Goffaux and Dakin (2010) we expected human observers to show horizontally-peaked masking, similar to the ideal observer, for upright but not inverted face stimuli.We also examined whether individual differences in face identification accuracy and/or the face inversion effect (Bruce *et al.*, 1999; Konar *et al.*, 2010; Sekuler *et al.*, 2004) can be explained by differences in the use of horizontal information. Specifically, if the preferential use of horizontal information is associated with face identification, then there ought to be a positive correlation between face identification accuracy and the strength of horizontal tuning.

2.2 Materials and methods

2.2.1 Observers

Thirty-two observers (9 male, 23 female; average age 21 years) participated in the experiment. All observers were naïve to the purpose of the experiment and had normal or corrected-to-normal Snellen acuity. Observers were paid $10/hour or given course credit for their participation. All experimental protocols were approved by the McMaster University Research Ethics Board, and informed consent was collected prior to initiation of the experiment.

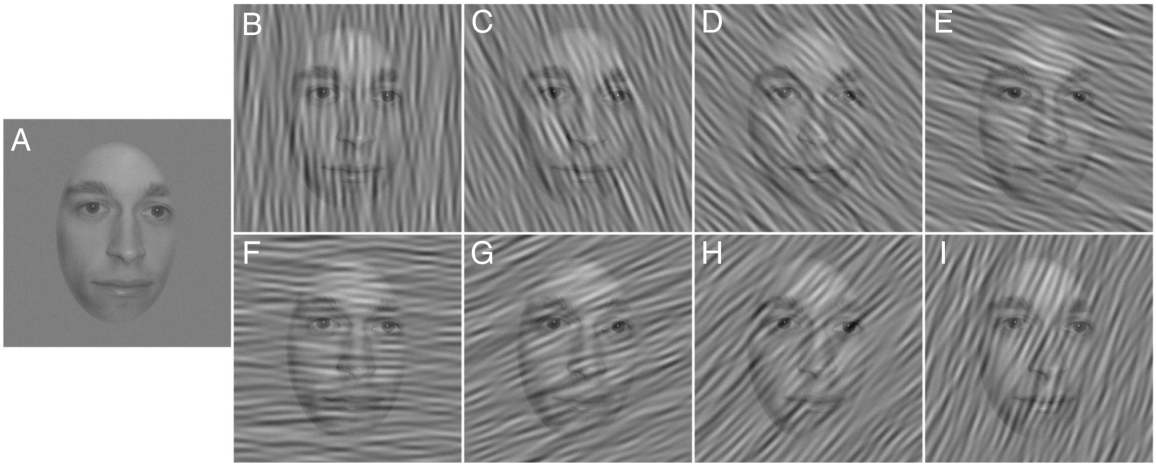


Figure 2.2: High-contrast examples of stimuli in each condition. (A) White noise only. (B-I) White noise and orientation filtered noise with centre orientations ranging from -90 deg (B) to 67.5 deg (I) in 22.5 degree steps.

2.2.2 Stimuli

Stimuli were generated on an Apple Macintosh G4 computer using MATLAB and the Psychophysics and Video Toolboxes (Brainard, 1997; Pelli, 1997). Stimuli were presented on a 21-inch Apple Studio display with a resolution of 1280 $\times$ 1024 pixels and

a frame rate of 85 Hz. Average luminance, which was 30.8 cd/m^2, was held constant throughout the experiment. The face stimuli were based on digitized photographs of 5 male and 5 female models (average age 24 years) with no visible piercings, facial hair, or eye glasses. Models were photographed as they turned their head to face a variety of gaze directions, each separated by 4.5 degrees of visual angle. In this way, each identity was represented by a variety of images with viewpoints to the left and the right, as well as one frontal view. Each photograph was cropped to remove external features such as hair, ears, and chin. The faces were centred in a 372 $\times$ 372 pixel matrix which subtended $4.6^\circ \times 4.6^\circ$ at the viewing distance of 60 cm. See Gaspar *et al.* (2008a) for more details about the stimuli.

Two independent Gaussian noise fields were added to the stimulus on every trial. One was an unfiltered (i.e., white) Gaussian noise with an RMS contrast of 0.028. The other noise field was filtered with an ideal, band-pass orientation filter with a full bandwidth of 23 degrees centered at one of eight orientations ranging from -90 deg (vertical), through 0 deg (horizontal), to 67.5 deg in 22.5 deg steps. The RMS contrast of the filtered noise was 0.14 prior to filtering. In a ninth condition, the contrast of the filtered noise was set to zero, so that the stimuli were embedded only in unfiltered, white Gaussian noise. Figure 2.2 demonstrates a stimulus masked in each of the different orientation conditions.

2.2.3 Procedure

Participants viewed the display binocularly, and a chin/head rest was used to stabilize the viewing position. Each trial began with a small, high-contrast fixation point presented at the centre of the screen for 500 ms. The fixation point was extinguished

and, after a delay of 200 ms, a face embedded in noise was presented for 250 ms. On each trial, a random viewpoint was selected for the current identity to discourage the use of simple image-matching strategies. Following the stimulus, a response selection screen containing noise-free, high-contrast (RMS contrast = 0.3) frontal views of the 10 face identities was presented and the observer selected the target with a mouse click. Note that frontal views only appeared on the response selection screen; target stimuli always were presented with viewpoints to the left or the right. Feedback was provided in the form of 600 and 200 Hz tones following correct and incorrect responses, respectively.

2.2.4 Design

Observers completed the experiment over the course of two sessions, separated by approximately 24 hours. Within each session, observers completed two blocks of trials: one block used upright stimuli, and the other used inverted stimuli. The order of face orientation blocks was counter-balanced across observers. The orientation of the faces in the response selection screen was the same as the orientation of the target stimuli. Noise conditions (eight orientation-filtered noises plus one white noise) were intermixed randomly within blocks. Face RMS contrast was varied across trials with the FAST toolbox, a Bayesian adaptive threshold estimator (Vul and MacLeod, 2010).

Two thresholds per noise condition were measured simultaneously within each block. A block ended when the threshold estimates for each condition were based on at least 20 trials and had a 95% confidence interval of less than 0.3 log units. Prior to the experiment, we were uncertain whether the strength of masking produced by oriented masking noise depended on the response accuracy used to define threshold.

Therefore, we measured thresholds using two criterion levels of response accuracy: For 16 subjects, threshold was defined as the RMS contrast needed to achieve 67% correct responses, and for the remaining 16 subjects threshold was defined as the RMS contrast needed to achieve 50% correct. In the following sections we refer to these two groups as the t_{67} and t_{50} groups.

2.2.5 Data analysis

The two thresholds for each condition in each session were averaged to form a single dependent measure. In each block, orientation masking was defined as the threshold obtained with an orientation-filtered noise divided by the threshold in the unfiltered (i.e., white) noise condition. These masking ratios were log transformed prior to analysis.

2.2.6 Ideal observer analysis

The ideal observer is an optimal decision maker that achieves the best possible performance on a task given the available stimulus information (Bennett and Banks, 1987; Banks *et al.*, 1991; Tjan *et al.*, 1995). If information within a given orientation band is not relevant for the task, then the performance of the ideal observer should not vary when that orientation band is masked. On the other hand, if information in a given orientation band is critical for the task, then performance of the ideal observer should be impaired when that orientation band is masked.

For a task like ours that uses white noise, the ideal observer is a cross-correlator that measures the *a posteriori* probability of each stimulus identity given a particular noisy input (Tjan *et al.*, 1995). If R is the noisy stimulus, σ^2 is the variance of the

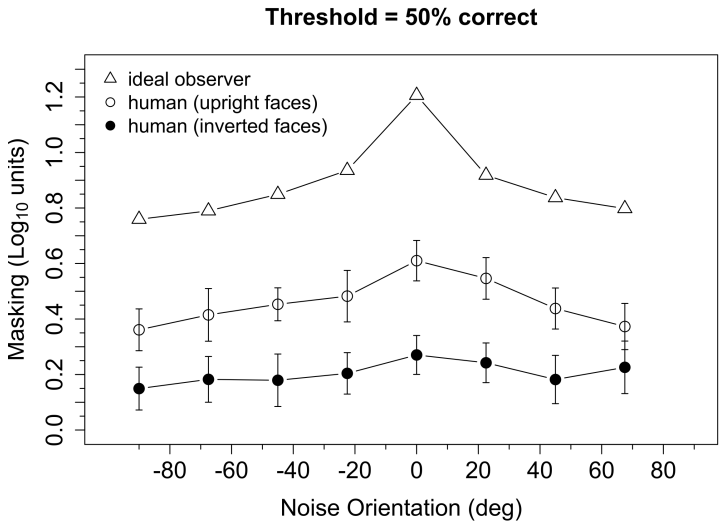
noise, T_{ij} is the jth view of the ith identity, and $P(T_{ij})$ is the *a priori* probability of being shown T_{ij}, then the the ideal observer selects the identity i that maximizes the function

$$\sum_{j} exp\left(-\frac{1}{2\sigma^2}\left\|R - T_{ij}\right\|^2\right) P(T_{ij}) \tag{2.1}$$

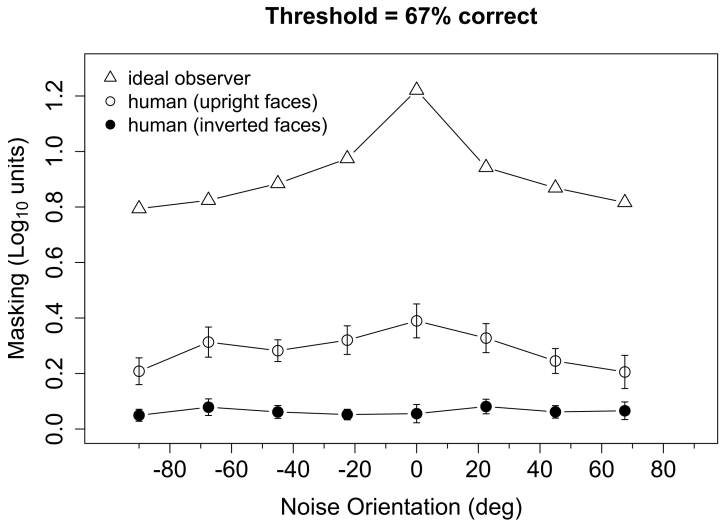
where $\|R - T_{ij}\|^2$ is defined as the Euclidian distance between the image and the template, and is equivalent to maximizing the cross correlation RT_{ij} between the stimulus and template when all the templates contain the same energy (Tjan *et al.*, 1995).

Our experiments used filtered noise, and therefore the ideal observer used templates that were adjusted to take into account the fact that noise power varies as a function of orientation. This adjustment can be carried out by computing the product, in the Fourier domain, of the original template and a pre-whitening filter that removes the noise correlations in the stimulus (Myers *et al.*, 1985; Eckstein *et al.*, 1997). These adjusted templates were used to maximize Equation 2.1.

We used computer simulations to calculate the performance of the ideal observer on our task. The stimuli, procedure, and design were identical to those used in the main experiment with the exception that we did not include an inverted condition because the ideal observer's performance is identical for upright and inverted faces. We simulated 10 sessions, yielding a total of 20 thresholds per condition. The mean of the 20 thresholds in each condition was calculated and utilized for all subsequent analyses.



(a)



(b)

Figure 2.3: Log transformed masking ratios plotted as a function of noise orientation for the ideal observer and human observers with upright and inverted face stimuli. Threshold was defined as the RMS contrast needed to achieve 50% or 67% correct responses (top and bottom panels, respectively), and the masking ratio was defined as the mean of the log-transformed ratios of masked to unmasked contrast thresholds. Error bars represent ±1 SEM.

2.3 Results

All statistical analyses were performed with R (R Core Team, 2015). The Huynh-Feldt correction, $\tilde{\epsilon}$, was used to adjust p values of F tests conducted with within-subject variables to correct for violations of sphericity (Maxwell and Delaney, 2004).

2.3.1 Ideal observer

Figure 2.3 plots log-transformed masking ratios as a function of noise orientation for the ideal and human observers. Consider first the ideal observer. When a particular orientation is masked by filtered noise, the ideal observer is forced to rely more heavily on information carried in the other orientation bands. Hence, the amount of masking obtained for each noise orientation is related to the amount of face identification information carried in each orientation band. Obtaining no masking would indicate that no information is carried in that orientation band, whereas a large masking ratio indicates that significant information is carried in that band. The ideal masking ratios in Figure 2.3 indicate that the amount of identification information was greatest for orientations near 0 deg i.e. horizontal), least for orientations near $\pm$ 90 deg (vertical), and intermediate for orientations near $\pm$ 22.5, $\pm$ 45, and $\pm$ 67 deg. This result highlights the fact that different orientations do, in fact, carry different amounts of physical information for our face discrimination task, as previously suggested by Dakin and Watt (2009). Specifically, there is more information for identification in the physical stimulus around horizontal orientations than around vertical orientations.

2.3.2 Human observers

Masking obtained from human observers as a function of noise orientation with upright and inverted faces is shown in Figure 2.3. A preliminary analysis indicated that log-transformed masking ratios, averaged across noise orientations, were greater than zero at both threshold criteria and both face orientations (Table 2.1). However, inspection of Figure 2.3 suggests that masking was greater in the t_{50} group than in the t_{67} group, and greater for upright faces than inverted faces. Furthermore, masking obtained with upright faces appeared to vary systematically with noise orientation, but was nearly independent of noise orientation with inverted faces. These observations were confirmed by a 2 (threshold criteria) $\times$ 2 (face orientation) $\times$ 8 (noise orientation) ANOVA performed on log-transformed masking ratios: The main effects of threshold criteria ($F(1, 30) = 6.38$, $p = 0.017$), face orientation ($F(1, 30) = 20.54$, $p < 0.001$), and noise orientation ($F(7, 210) = 8.28$, $p < 0.001$) were significant, as was the interaction between face and noise orientation ($F(7, 210) = 4.51$, $p = 0.001$).

Group	Face Orientation	M	CI_{95}
t_{50}	upright	0.418	[0.307, 0.613]
t_{50}	inverted	0.204	[0.045, 0.364]
t_{67}	upright	0.286	[0.187, 0.385]
t_{67}	inverted	0.063	[0.017, 0.109]

Table 2.1: Mean and 95% Confidence Intervals for masking averaged across noise orientations.

The significant interaction between face and noise orientation reflects the fact the masking functions obtained with upright faces, but not inverted faces, exhibited a peak near 0 deg. To quantify this interaction, we computed a measure of orientation

tuning for each subject by estimating the slope of a regression line that related masking to the orientation of the noise. Initially, we fit two regression lines to the masking data: one to the ascending part of the masking function for noise orientations from -90 to 0 deg, and another to the descending part of the function (i.e., noise orientations from 0 to 90 deg). However, the slopes of the ascending and descending parts of the curve were significantly correlated (upright: $r = 0.81$, $t(30) = 7.58$, $p < 0.001$; inverted: $r = 0.69$, $t(30) = 5.26$, $p < 0.001$). Furthermore, a 2 (face orientation) x 2 (threshold criterion) x 2 (masking function part: ascending vs. descending) ANOVA on the slopes found that the main effect of masking function part ($F(1, 30) = 2.82$, $p = 0.10$), as well as all of the interactions with that factor ($F \leq 2.93$ & $p \geq 0.10$, in each case), were not significant. Therefore, to simplify our analyses, we averaged the two measures of masking obtained with $\pm$22.5, $\pm$45, and $\pm$67.5 deg noise, computed a regression line for masking at noise orientations of -90, $\pm$67.5, $\pm$45, $\pm$22.5, and 0 deg, and used the slope of the regression line as the index of the horizontal tuning of masking. Boxplots of slopes of the regression lines are shown in Figure 2.4: tuning appeared to be significantly higher for upright faces than inverted faces, and slightly higher in the t_{50} group than the t_{67} group. A 2 (face orientation) $\times$ 2 (threshold criterion) ANOVA on the horizontal tuning measures confirmed these observations: the main effects of face orientation ($F(1, 30) = 49.12$, $p < 0.001$) and threshold criterion ($F(1, 30) = 4.84$, $p = 0.035$) were significant, but the face orientation $\times$ threshold criterion interaction was not significant ($F(1, 30) = 0.33$, $p = 0.57$). One subject in the t_{67} group had an unusually low slope in the inverted face condition (Figure 2.4). When this subject was removed from the analysis, the main effect of face orientation was significant ($F(1, 29) = 45.42$, $p < 0.001$) but the main effect of threshold criterion

($F(1, 29) = 3.60$, $p = 0.07$) and interaction between face orientation and threshold criterion ($F(1, 29) = 0.27$, $p = 0.60$) were not. Note that nearly identical results were obtained from separate ANOVAs that were conducted on tuning measures derived from the ascending and descending parts of the masking function. Separate t tests indicated that tuning differed significantly from zero in all conditions ($t(15) \geq 2.88$, $p \leq 0.011$, in each case) except for inverted faces in the t_{67} group with ($t(15) = 0.12$, $p = 0.91$) or without ($t(14) = 1.68$, $p = 0.12$) the outlier. Together, these results suggest that the orientation selectivity of masking was greater for upright faces than inverted faces.

2.3.3 Absolute efficiency

Our results demonstrate that human observers used information across a wide range of orientations for upright face discrimination, with more weight being given to orientations closer to horizontal. However, as our ideal observer analysis showed, more information is available in the horizontal orientations for face discrimination. To what extent does the human pattern of results simply reflect the variation of information across orientations? We addressed this question by calculating absolute efficiency of human observers as a function of orientation for upright and inverted faces. In an identification task such as ours, absolute efficiency is defined as the squared ratio of the ideal to human RMS contrast thresholds.

If human observers extracted information from all bands of orientation with equal efficiency, then absolute efficiency for faces embedded in filtered noise ought to be constant as a function of noise orientation. On the other hand, if human observers use information at a particular orientation relatively poorly, then masking noise at

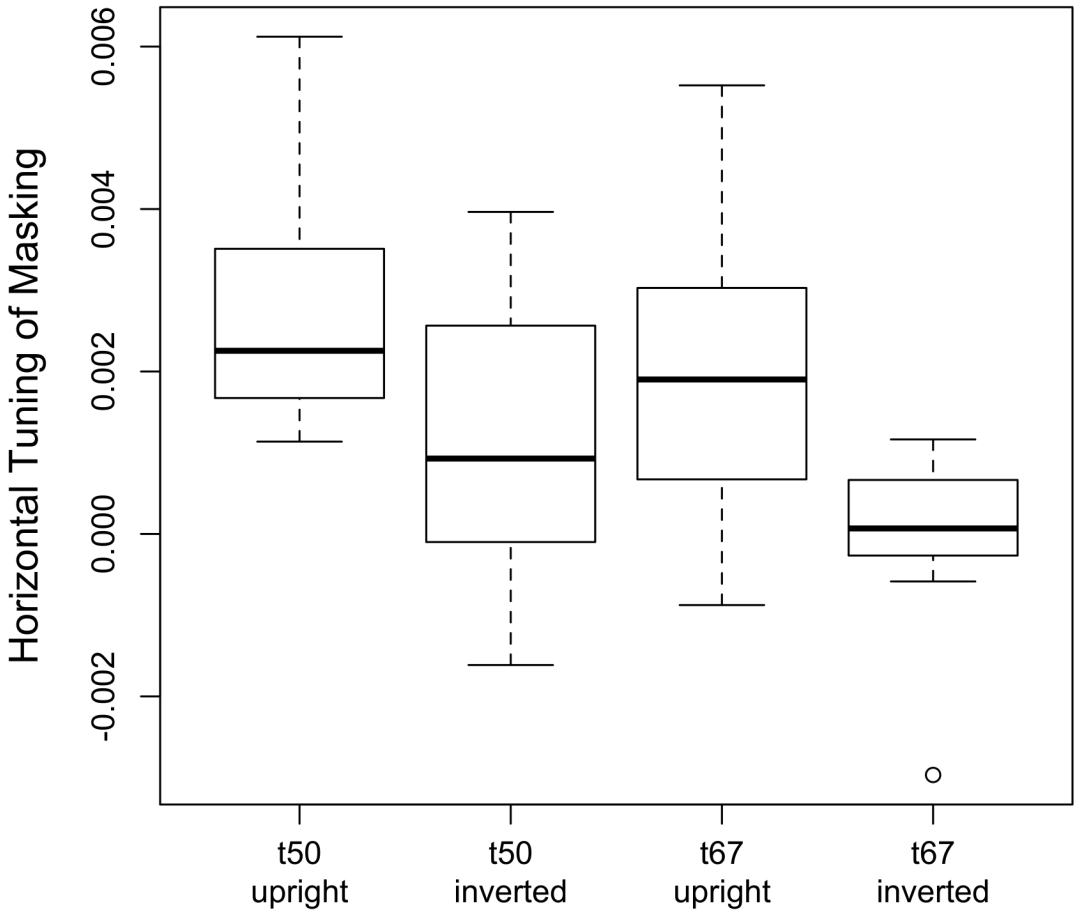


Figure 2.4: Orientation tuning of masking for upright and inverted faces in the t_{50} and t_{67} groups. Tuning was defined as the slope of the regression line fit to masking obtained with noise orientations of -90, 67.5, 45, 22.5, and 0 deg. Masking values at 67.5, 45, and 22.5 were defined as the average level of masking obtained at, respectively, $\pm$67.5, $\pm$45, and $\pm$22.5 deg. The horizontal line in each boxplot indicates the median; the upper and lower edges of each box indicate the 75th and 25th percentile, respectively.

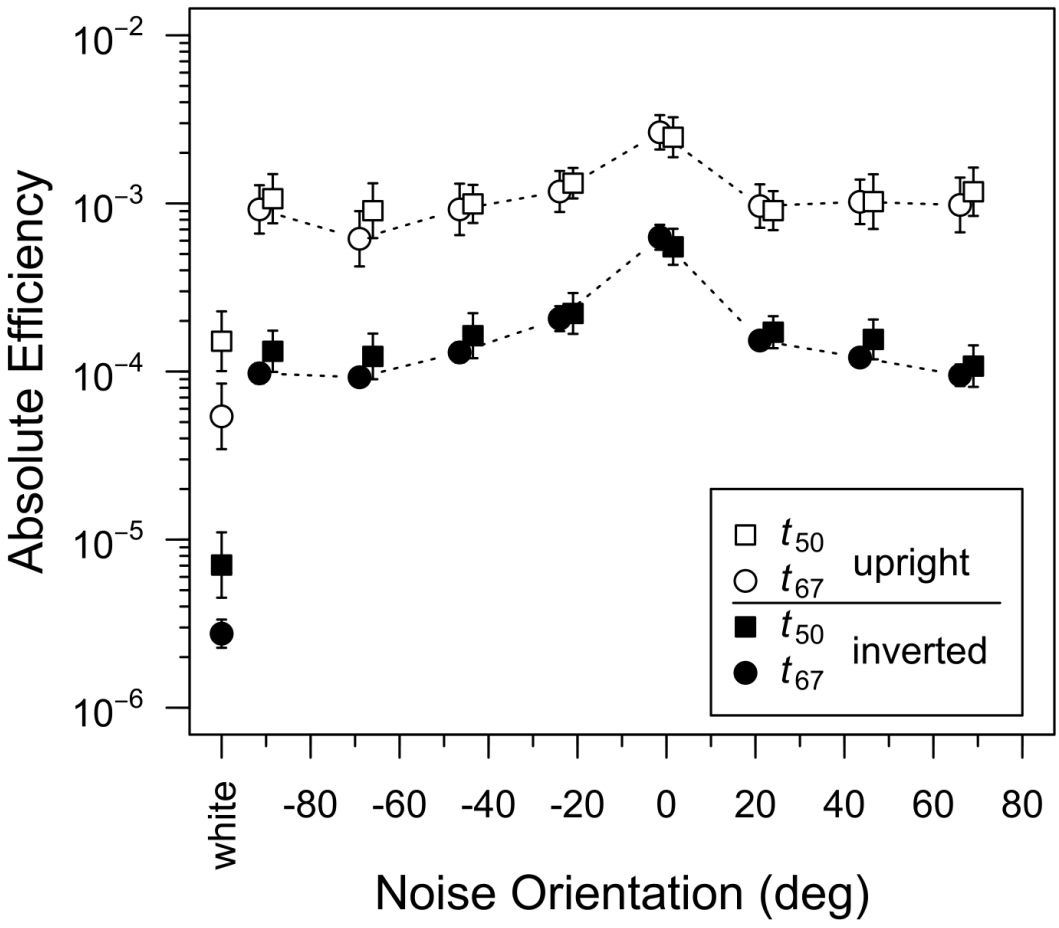


Figure 2.5: Absolute efficiency measured in the t_{50} and t_{67} groups plotted as a function of noise orientation for upright and inverted faces. The leftmost symbols represent efficiency in the white noise condition. Absolute efficiency is defined as the squared ratio of ideal and human RMS contrast thresholds. Points for the t_{50} and t_{67} groups have been offset slightly for clarity. Error bars, where visible, represent ± 1 SEM.

that orientation should increase threshold more in the ideal observer than in human observers and therefore result in *higher* efficiency. Figure 2.5 plots absolute efficiency as a function of noise orientation for upright and inverted faces. Consistent with previous reports, efficiency obtained with a white noise mask was higher for upright than inverted faces (Gaspar *et al.*, 2008b), although efficiency was quite low at both face orientations (Gold *et al.*, 1999, 2004). A 2 (threshold criterion) $\times$ 2 (face orientation) ANOVA on log-transformed efficiency in the white noise conditions yielded significant main effects of threshold criterion ($F(1, 30) = 4.38$, $p = 0.045$) and face orientation ($F(1, 30) = 110.6$, $p < 0.001$); the interaction between threshold criterion and noise orientation was not significant ($F(1, 30) = 0.024$, $p = 0.87$).

With both upright and inverted faces, average efficiency was higher in conditions that used orientation-filtered noise than in the white noise condition, which is due to the fact that the addition of filtered noise increased thresholds less (i.e., masking was lower) in human observers than the ideal observer (Figure 2.3). As discussed previously, this result suggests that human observers used information at all orientations, albeit less efficiently than the ideal observer. Furthermore, the fact that the addition of filtered noise increased efficiency more for inverted faces than upright faces is consistent with the observation that efficiency is lower overall for inverted faces. Efficiency also varied with noise orientation, although the variation appeared greater with inverted faces. Adding filtered noise eliminated the difference between efficiency in the t_{50} and t_{67} groups. These observations were confirmed with a 2 (threshold criterion) $\times$ 2 (face orientation) $\times$ 8 (noise orientation) ANOVA on log-transformed efficiency: The main effect of threshold criterion was not significant ($F(1, 30) = 0.21$, $p = 0.65$), but the main effects of face orientation ($F(1, 30) = 95.13$, $p < 0.001$)

and noise orientation ($F(7, 210) = 41.57$, $\tilde{\epsilon} = 0.748$, $p < 0.001$) were significant, as was the interaction between face and noise orientation ($F(7, 210) = 4.51$, $\tilde{\epsilon} = 0.939$, $p < 0.001$). The significant interaction reflected the fact that the difference between efficiency for upright and inverted faces was ≈ 1 log unit when the noise orientation was -90 and ± 67.5 deg but only ≈ 0.6 log units with 0 deg noise. Again, because suboptimal use of a particular orientation band should result in *higher* efficiency, this interaction suggests that human observers were suboptimal in their use of horizontal information, particularly so with inverted faces. Follow up analyses indicated that the effect of noise orientation was significant for both upright ($F(7, 210) = 13.38$, $\tilde{\epsilon} = 0.826$, $p < 0.001$) and inverted faces ($F(7, 210) = 39.14$, $\tilde{\epsilon} = 0.793$, $p < 0.001$), suggesting that observers differ quantitatively in their use of orientation information following inversion, relying on horizontal contours in both cases, but less effectively with inverted faces.

Absolute efficiency was lower in the t_{67} group than the t_{50} group in the white noise condition but not the oriented noise conditions. The group difference in the white noise condition is consistent with the hypothesis that the psychometric function relating face contrast to response accuracy in human observers was shallower than the psychometric function for the ideal observer: contrast had to be increased more in human observers to increase accuracy from 50% to 67% correct. Conversely, the lack of a group difference in the oriented noise conditions implies that the psychometric functions for human and ideal observers had similar slopes in those conditions. However, it is unclear why the slope of the psychometric function differed in the white and oriented noise conditions.

	df	Sum Sq	Mean Sq	F value	Pr(>F)
Upright Faces:					
criterion	1	0.91	0.91	7.80	0.009
tuning	1	0.90	0.90	7.76	0.009
criterion × tuning	1	0.01	0.01	0.09	0.762
Residuals	28	3.26	0.12		
Inverted Faces:					
criterion	1	0.81	0.81	8.48	0.007
tuning	1	0.02	0.02	0.26	0.616
criterion × tuning	1	0.00	0.00	0.00	0.985
Residuals	28	2.67	0.10		

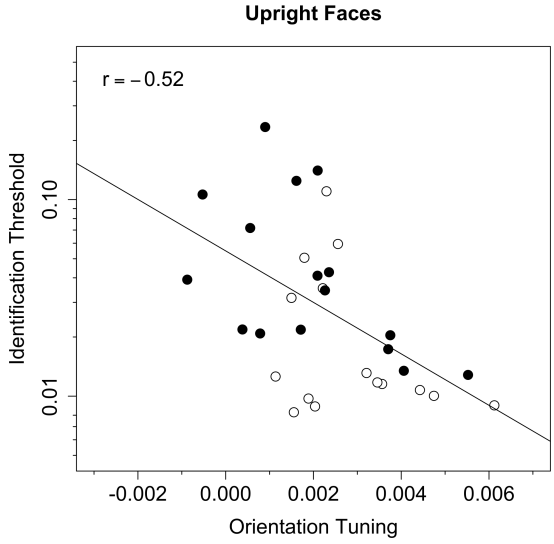
Table 2.2: ANOVA tables for linear models predicting identification threshold in the white-noise condition with upright faces (top) and inverted faces (bottom). The variable criterion refers to the criterion used to define threshold (i.e., the t_{67} and t_{50} groups); tuning refers to the orientation tuning of masking (see Figure 2.4). Type-I tests (i.e., sequential sums-of-squares) are shown: each row evaluates the change in the residual sum of squares that is produced by adding the predictor variable to a model that contains all of the predictor variables listed above it.

2.3.4 Correlation analysis

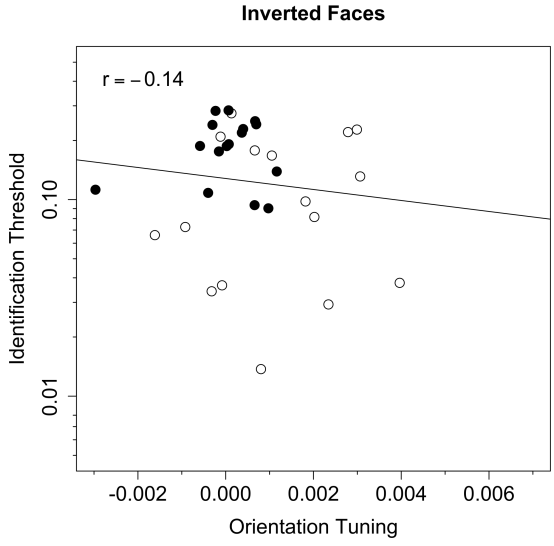
To assess the association between orientation tuning and face identification threshold, we evaluated linear models that included log-transformed threshold in the white noise condition as the dependent variable, and threshold criterion (i.e., t_{67} *vs.* t_{50}), orientation tuning (see Figure 2.4), and the interaction between criterion and tuning as predictor variables. These models allowed us to quantify the variance in identification thresholds that could be explained by variance in each of the predictor variables. Note that the derivation of orientation tuning did not include thresholds in the white noise condition, and therefore the two variables were not necessarily related. Models for upright and inverted faces were evaluated separately (see Table

2.2). As expected, the effect of threshold criterion was significant for both upright and inverted faces: thresholds in the white noise condition were higher in the t_{67} group than the t_{50} group. After statistically controlling for the effect of criterion, the effect of orientation tuning was significant for upright but not inverted faces, and the interaction between criterion and tuning was not significant for either face orientation. These results indicate that tuning was correlated with identification thresholds for upright face but not inverted faces, and that that the correlation did not differ between the t_{67} and t_{50} groups. After combining the t_{67} and t_{50} groups, the Pearson correlation between log-transformed threshold and orientation tuning was -0.52 ($CI_{95} = [-0.73, -0.21]$, $t(30) = -3.33$, $p = 0.0023$) for upright faces and -0.12 ($CI_{95} = [-0.45, 0.24]$, $t(30) = -0.65$, $p = 0.52$) for inverted faces (Figure 2.6). Hence, greater orientation tuning was associated with lower identification thresholds for upright faces, but not inverted faces, in the white noise condition. Essentially the same results were obtained when the data were re-analyzed after removing the unusually low tuning score obtained with inverted faces from one subject in the t_{67} group (see Figure 2.4).

Given that orientation tuning predicted identification threshold for upright, but not inverted faces, it follows that orientation tuning may predict the size of the face inversion effect, defined here as the difference between the log transformed contrast thresholds for upright and inverted faces. To assess the association between the face inversion effect and orientation tuning, we evaluated a linear model that used the face inversion effect as the dependent variable and threshold criterion, orientation tuning for upright faces, and the criterion $\times$ tuning interaction as predictor variables. The effect of orientation tuning was significant, ($F(1, 28) = 9.14$, $p = 0.005$), but the



(a)



(b)

Figure 2.6: Identification threshold plotted against orientation tuning for upright and inverted faces. Data from the t_{67} and t_{50} groups are represented by the filled and open symbols, respectively. The dotted line represents the best-fitting (least-squares) line fit to the data from both groups. The Pearson correlation between identification threshold and orientation tuning was significant for upright ($r = -0.52$) but not inverted ($r = -0.14$) faces.

effects of threshold criterion ($F(1, 28) = 0.03$, $p = 0.86$) and the criterion $\times$ tuning interaction ($F(1, 28) = 0.21$, $p = 0.65$) were not. Because the effects of threshold criterion and the interaction were not significant, those two predictor variables were dropped from the model, and the best-fitting line relating upright face orientation tuning to the face inversion effect was computed: the face inversion effect was positively correlated with upright tuning ($r = 0.48$, $\text{CI}_{95} = [0.16, 0.71]$, $t(30) = 3.03$, $p = 0.005$). Hence, greater orientation tuning for upright faces was associated with a larger face inversion effect (see Figure 2.7). A model using orientation tuning for inverted faces as a predictor variable fit the data poorly ($R^2 = 0.015$, $F(3, 28) = 0.14$, $p = 0.933$), and the face inversion effect was not associated with any of the predictor variables ($F < 1$ and $p > 0.5$ in all cases).

2.4 Discussion

In this experiment we found that human observers preferentially use horizontal information to identify upright more than inverted faces. This result is reflected in the masking functions of our human observers, as well as the significant linear regression between masking and noise orientation with upright faces. Moreover, our ideal observer analysis also obtained the strongest masking with horizontal noise, which suggests that more information relevant to face identification is carried in this band. Therefore, our findings suggest that human observers exploit diagnostic orientation information more efficiently when identifying upright faces. These results are consistent with previous demonstrations of a preference for horizontal information in upright but not inverted face discrimination (Goffaux and Dakin, 2010), and the presence of structured bands of horizontal information in face stimuli (Dakin and

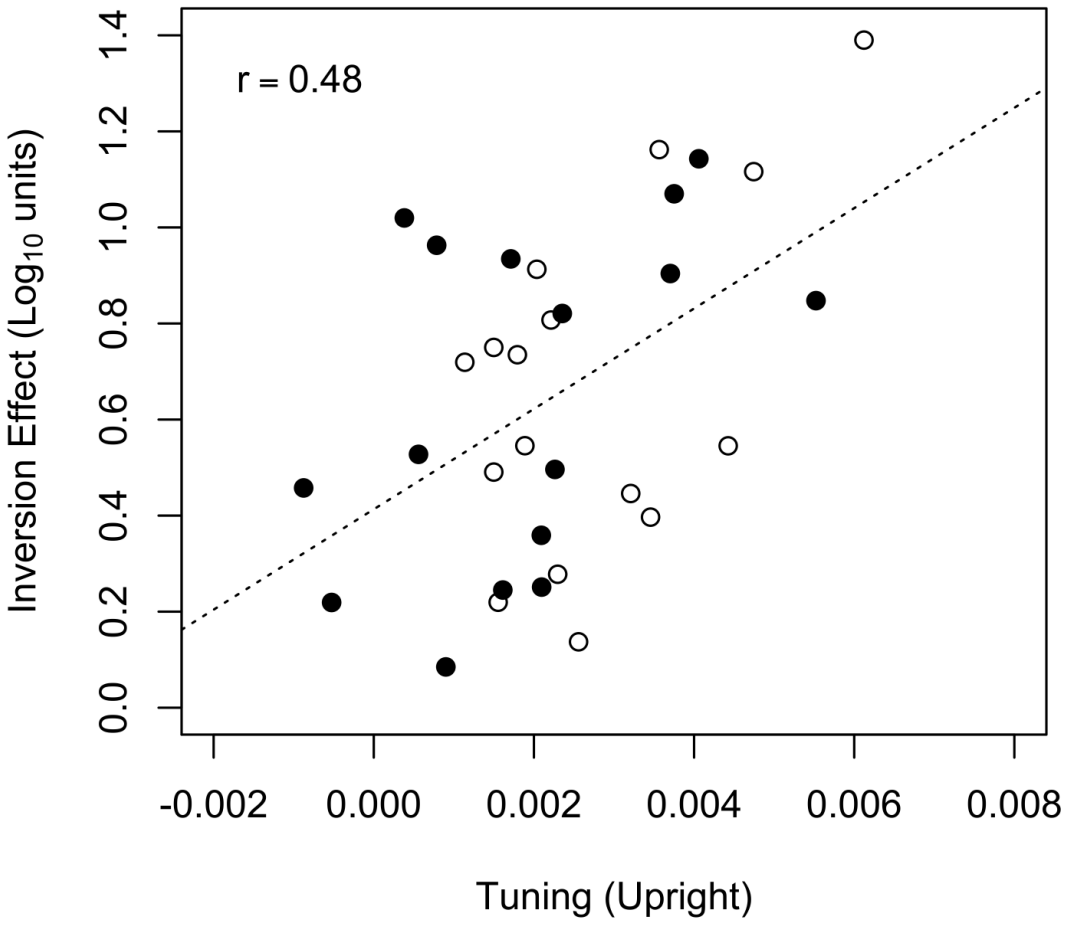


Figure 2.7: The face inversion effect plotted against orientation tuning for upright faces. Data from the t_{67} and t_{50} groups are represented by the filled and open symbols, respectively. The dotted line represents the best-fitting (least-squares) line fit to the data from both groups. The Pearson correlation between the face inversion effect and upright tuning ($r = 0.48$) was significant.

Watt, 2009). Our ideal observer analysis showed that information relevant to identification is carried at every orientation, with relatively more information available in the horizontal band. In fact, although our human observers demonstrated significant horizontal orientation tuning, particularly for upright faces, our absolute efficiency results indicate that, compared to other orientations, they were suboptimal in their use of the extensive information carried in the horizontal band. However, this failure to take full advantage of the additional horizontal information is less pronounced with upright than inverted faces. Given these results, it follows that observers who best utilize horizontal information should also demonstrate the best overall face identification performance. Our results were consistent with this hypothesis: We found a significant negative correlation between orientation tuning and identification thresholds for upright faces but not inverted faces. We also found a significant correlation between orientation tuning for upright faces and the size of the face inversion effect.

Together, these results are consistent with recent demonstrations that the key difference between upright and inverted face processing is the manner in which observers encode horizontal structure (Goffaux and Dakin, 2010; Goffaux *et al.*, 2011). However, we have demonstrated that face stimuli do indeed carry more diagnostic information in the horizontal band, and that differential sensitivity to this information explains much of the variance in upright face identification and the face inversion effect. Moreover, we find that observers are sensitive to information in the horizontal band when processing upright and inverted faces, but use this information less effectively with inverted faces. As such, the current results are consistent with the idea that the face inversion effect reflects quantitative differences in the efficiency with which observers extract diagnostic information from upright and inverted faces

(Riesenhuber *et al.*, 2004; Sekuler *et al.*, 2004; Yovel and Kanwisher, 2004). Indeed, previous results using noise masking techniques have demonstrated only subtle differences in spatial sampling following inversion (Sekuler *et al.*, 2004) or perceptual learning (Gold *et al.*, 2004), coupled with changes in calculation efficiency (Gaspar *et al.*, 2008b).

A great deal of information for face identification is clustered around the eye and eyebrow region (Gold *et al.*, 1999; Sadr *et al.*, 2003; Gold *et al.*, 2004; Sekuler *et al.*, 2004; Vinette *et al.*, 2004; Gaspar *et al.*, 2008b; Keil, 2009), and these regions are rich in horizontal structure (Dakin and Watt, 2009). Human observers likely learn to efficiently extract diagnostic information as they become experts with upright faces throughout their development (de Heering *et al.*, 2012). Therefore, although horizontal information appears to be critical for upright face identification, in other tasks such as emotion discrimination or gender discrimination, different regions of the face or different orientations may be critical (Smith *et al.*, 2005). It remains unclear whether orientation tuning is associated with behavioural performance in these tasks. Some aspects of face perception appear to differ across culture (Jack *et al.*, 2009), age (Carey *et al.*, 1980; Bruce *et al.*, 2000; Mondloch *et al.*, 2002; Boutet and Faubert, 2006; Habak *et al.*, 2008; Rousselet *et al.*, 2009; Obermeyer *et al.*, 2012), specialized subject populations (Langdell, 1978; Archer *et al.*, 1992; Duchaine and Nakayama, 2005), and contrast polarity (Vuong *et al.*, 2005; Russell *et al.*, 2006; Gaspar *et al.*, 2008b), and more work is needed to elucidate how orientation tuning may be associated with these phenomena. It also remains unclear whether orientation tuning can be modulated with perceptual learning. However, if orientation tuning is impaired in populations with impairments in face perception, and orientation tuning can be

modulated with perceptual learning, then this line of research may prove fruitful in developing focused training programs to help ameliorate the deficits experienced by these individuals.

Bibliography

Archer, J., Hay, D. C., and Young, A. W. (1992). Face processing in psychiatric conditions. *British Journal of Clinical Psychology*, **31**(1), 45–61.

Banks, M. S., Sekuler, A. B., and Anderson, S. J. (1991). Peripheral spatial vision: limits imposed by optics, photoreceptors, and receptor pooling. *Journal of the Optical Society of America. A, Optics and Image Science*, **8**, 1775–1787.

Bennett, P. J. and Banks, M. S. (1987). Sensitivity loss in odd-symmetric mechanisms and phase anomalies in peripheral vision. *Nature*, **326**, 873–876.

Boutet, I. and Faubert, J. (2006). Recognition of faces and complex objects in younger and older adults. *Memory & cognition*, **34**(4), 854–64.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**(4), 433–6.

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J. B., Burton, A. M., and Miller, P. (1999). Vertification of Face Identities from Images Captured on Video. *Journal of Experimental Psychology: Applied*, **5**(4), 339–360.

Bruce, V., Campbell, R. N., Doherty-Sneddon, G., Import, A., Langton, S., Mcauley, S., and Wright, R. (2000). Testing face processing skills in children. *British Journal of Development Psychology*, **18**, 319–333.

Carey, S., Diamond, R., and Woods, B. (1980). Development of Face Recognition A Maturational Component ? *Developmental Psychology*, **16**(4), 257–269.

Dakin, S. C. and Watt, R. J. (2009). Biological bar codes in human faces. *Journal of Vision*, **9**(4), 1–10.

de Heering, A., Aljuhanay, A., Rossion, B., and Pascalis, O. (2012). Early deafness increases the face inversion effect but does not modulate the composite face effect. *Frontiers in psychology*, **3**(April), 124.

Diamond, R. and Carey, S. (1986). Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology: General*, **115**(2), 107–17.

Duchaine, B. C. and Nakayama, K. (2005). Dissociations of face and object recognition in developmental prosopagnosia. *Journal of Cognitive Neuroscience*, **17**(2), 249–61.

Eckstein, M. P., Ahumada, A. J., and Watson, A. B. (1997). Visual signal detection in structured backgrounds. II. Effects of contrast gain control, background variations, and white noise. *Journal of the Optical Society of America. A, Optics and Image Science*, **14**(9), 2406–19.

Farah, M. J., Tanaka, J. W., and Drain, H. M. (1995). What Causes the Face Inversion Effect? *Journal of Experimental Psychology: Human Perception and Performance*, **21**(3), 628–634.

Gaspar, C. M., Sekuler, A. B., and Bennett, P. J. (2008a). Spatial frequency tuning of upright and inverted face identification. *Vision Research*, **48**(28), 2817–26.

Gaspar, C. M., Bennett, P. J., and Sekuler, A. B. (2008b). The effects of face inversion and contrast-reversal on efficiency and internal noise. *Vision Research*, **48**(8), 1084–95.

Goffaux, V. and Dakin, S. C. (2010). Horizontal information drives the behavioral signatures of face processing. *Frontiers in Perception*, **1**, 1–14.

Goffaux, V., van Zon, J., and Schiltz, C. (2011). The horizontal tuning of face perception relies on the processing of intermediate and high spatial frequencies. *Journal of Vision*, **11**, 1–9.

Gold, J. M., Bennett, P. J., and Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research*, **39**(21), 3537–60.

Gold, J. M., Sekuler, A. B., and Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science*, **28**(2), 167–207.

Habak, C., Wilkinson, F., and Wilson, H. R. (2008). Aging disrupts the neural transformations that link facial identity across views. *Vision research*, **48**(1), 9–15.

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

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., and Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology*, **19**(18), 1543–8.

Keil, M. S. (2009). "I look in your eyes, honey": internal face features induce spatial frequency preference for human face processing. *PLoS Computational Bology*, **5**(3).

Konar, Y., Bennett, P. J., and Sekuler, A. B. (2010). Holistic Processing Is Not Correlated With Face-Identification Accuracy. *Psychological Science*, **21**(1), 38–43.

Langdell, T. (1978). Recognition of faces: an approach to the study of autism. *Journal of Child Psychology and Psychiatry*, **19**(3), 255–68.

Maxwell, S. E. and Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective*, volume 1. Psychology Press.

Mondloch, C. J., Le Grand, R., and Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, **31**(5), 553–566.

Myers, K. J., Barrett, H. H., Borgstrom, M. C., Patton, D. D., and Seeley, G. W. (1985). Effect of noise correlation on detectability of disk signals in medical imaging. *Journal of the Optical Society of America. A, Optics and Image Science*, **2**(10), 1752–9.

Obermeyer, S., Kolling, T., Schaich, A., and Knopf, M. (2012). Differences between Old and Young Adults' Ability to Recognize Human Faces Underlie Processing of Horizontal Information. *Frontiers in aging neuroscience*, **4**(April), 3.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, **10**(4), 437–442.

R Core Team (2015). R: A Language and Environment for Statistical Computing.

Riesenhuber, M., Jarudi, I., Gilad, S., and Sinha, P. (2004). Face processing in humans is compatible with a simple shape-based model of vision. *Proceedings of the Royal Society B*, **271 Suppl**, S448–50.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta psychologica*, **128**(2), 274–89.

Rousselet, G. A., Husk, J. S., Pernet, C. R., Gaspar, C. M., Bennett, P. J., and Sekuler, A. B. (2009). Age-related delay in information accrual for faces: evidence from a parametric, single-trial EEG approach. *BMC neuroscience*, **10**, 114.

Russell, R., Sinha, P., Biederman, I., and Nederhouser, M. (2006). Is pigmentation important for face recognition? Evidence from contrast negation. *Perception*, **35**(6), 749–759.

Sadr, J., Jarudi, I., and Sinha, P. (2003). The role of eyebrows in face recognition. *Perception*, **32**(3), 285–293.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., and Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, **14**(5), 391–6.

Smith, M. L., Cottrell, G. W., Gosselin, F., and Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological science*, **16**(3), 184–9.

Tanaka, J. W. and Farah, M. J. (1993). Parts and Wholes in Face Recognition. *Journal of Experimental Psychology*, **46**(2), 225–245.

Tjan, B. S., Braje, W. L., Legge, G. E., and Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research*, **35**(21), 3053–69.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, **79**(4), 471–491.

Vinette, C., Gosselin, F., and Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: its in the eyes. *Cognitive Science*, **28**(2), 289–301.

Vul, E. and MacLeod, D. (2010). Functional Adaptive Sequential Testing (FAST a. 1.0).

Vuong, Q. C., Peissig, J. J., Harrison, M. C., and Tarr, M. J. (2005). The role of surface pigmentation for recognition revealed by contrast reversal in faces and Greebles. *Vision research*, **45**(10), 1213–23.

Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J. W., Bub, D. N., and Gosselin, F. (2010). Does face inversion change spatial frequency tuning? *Journal of experimental psychology. Human perception and performance*, **36**(1), 122–35.

Yin, R. K. (1969). Looking at Upside-Down Faces. *Journal of Experimental Psychology*, **81**(1), 141–145.

Young, A. W., Hellawell, D., and Hay, D. C. (1987). Configurational information in face perception. *Perception*, **16**(6), 747–59.

Yovel, G. and Kanwisher, N. (2004). Face perception: domain specific, not process specific. *Neuron*, **44**(5), 889–98.

Chapter 3

Effects of bandwidth on selectivity for horizontal structure during face identification

3.1 Introduction

A fundamental challenge in the study of face perception is to characterize the information extracted from face stimuli during a variety of tasks. In general, face identification is well characterized by use of a limited subset of the available information. For example, experiments using a variety of techniques have shown that observers focus primarily on the eyes when identifying faces (Gosselin and Schyns, 2001; Peterson and Eckstein, 2012; Sekuler *et al.*, 2004; Yarbus, 1967). Further, although diagnostic information is available at a variety of spatial scales, several studies suggest that face identification is driven by a limited range of spatial frequencies (Gaspar *et al.*, 2008b; Gold *et al.*, 1999; Näsänen, 1999; Willenbockel *et al.*, 2010). Such results may

be understood as the result of perceptual learning, whereby perceptual processing narrows as the visual system learns the most diagnostic information for the task at hand (Gauthier and Tarr, 1997; Heisz and Shore, 2008; LeGrand *et al.*, 2004; Tanaka, 2001).

The face inversion effect is a particularly compelling example of perceptual expertise. That face identification is severely impaired by picture-plane inversion has been known for many years (Valentine, 1988; Yin, 1969). This phenomenon is highly robust and appears across a variety of face-related tasks (Diamond and Carey, 1986; Farah *et al.*, 1998; Tanaka and Farah, 1993). Importantly, picture-plane inversion has no effect on the low-level information available to discriminate among identities. As such, any change in performance following face inversion must reflect differential processing on the part of the observer. One influential view is that inversion impairs the holistic processing of internal facial features, resulting in a qualitative shift of processing strategy (Rossion, 2008). Alternatively, the face inversion effect could simply result from qualitatively similar, but less efficient processing of the available information (Gaspar *et al.*, 2008a; Sekuler *et al.*, 2004). According to this framework, we should be able to identify a low level property of the stimulus that is used during both upright and inverted identification, but less efficiently in the latter case. To this end, several experiments have been conducted examining the effect of inversion on the spatial (Sekuler *et al.*, 2004) and spatial frequency (Gaspar *et al.*, 2008b; Willenbockel *et al.*, 2010) tuning of face identification, but no systematic effect of inversion on these measures has been found.

Recent evidence, however, points to another source of information related to face perception: the selective use of horizontal orientation structure. Dakin and Watt

(2009) first noted that horizontal structure in the Fourier domain is particularly diagnostic for face identification tasks, is unaffected by manipulations that do not commonly affect face identification (e.g., compression, pose variance), and is highly affected by manipulations that commonly impair face identification (e.g., contrast reversal, picture-plane inversion). This diagnostic structure also has been linked to the middle spatial frequency band preferred by human observers during face identification tasks (Goffaux *et al.*, 2011), and shown to be necessary for observation of a number of robust phenomena such as identity aftereffects and integrative processing (Goffaux and Dakin, 2010). In addition, the extent to which observers selectively utilize horizontal structure during an identification task is the first measure that has been able to predict both individuals' overall face identification ability and the size of their face inversion effect (Pachai *et al.*, 2013).

Although several studies now have demonstrated the criticality of observers' selective use of horizontal orientation structure for face identification, none has examined systematically the most appropriate bandwidth to measure this tuning. In their original study of orientation-filtered faces, Dakin and Watt (2009) used Gaussian-shaped filters with a standard deviation of 23 ° and subsequent studies followed suit (Huynh and Balas, 2014; Obermeyer *et al.*, 2012; Yu and Chung, 2011), with the exception of Goffaux and colleagues (Goffaux and Dakin, 2010; Goffaux *et al.*, 2011) who used a bandwidth of 14 ° to more closely match the orientation tuning properties of neurons in V1 (Ringach *et al.*, 2002). These studies also reported variations in human performance without evaluating how these variations could be anticipated as a result of the informational manipulation. Pachai *et al.* (2013) partially addressed these concerns using an ideal observer analysis to quantify the available orientation information.

This analysis revealed that the horizontal band contained the most diagnostic information, but, again, their filter bandwidth was a relatively large 22.5 °, providing only a coarse analysis of orientation tuning.

In this study, we sought to quantify more precisely the effects of filter bandwidth on horizontal tuning during face identification. To that end, faces were filtered to retain horizontal or vertical orientation structure, using ideal (sharp-edged) orientation filters and systematically varying bandwidth in 10 ° steps within-subjects. Further, to better measure the selective processing employed by human observers when identifying intact faces, we replaced the orientation components removed from the target face with components from the average of all 10 possible faces. This procedure ensured that the diagnostic band was unknown to the observer on a given trial and produced faces that appeared intact despite the filtering manipulation. Finally, we compared the results of our human observers to several simulated observers.

3.2 Methods

3.2.1 Observers

Thirty-two naïve observers (16 male, aged 19-33, $\mu = 22$ years) participated in the experiment. All observers had normal or corrected-to-normal Snellen acuity, and were paid $10/hour or given course credit for their participation. All experimental protocols were approved by the McMaster University Research Ethics Board, and informed consent was collected prior to the experiment.

3.2.2 Stimuli

Stimuli were generated using an Apple Macintosh G4 computer with MATLAB and the Psychophysics and Video Toolboxes (Brainard, 1997; Pelli, 1997), and were presented on a 21 inch Apple Studio display with a resolution of 1280 $\times$ 1024 and a frame rate of 85 Hz. Average luminance was held constant at 31 cd/m^2 throughout the experiment. The face images were generated using front-facing, digitized photographs of 5 male and 5 female models with no facial hair, eye glasses, or visible piercings. These photographs were cropped using a 198 $\times$ 140 pixel oval window, and were centred in a 256 $\times$ 256 matrix. At the viewing distance of 60 cm, these stimulus matrices subtended $7.6^\circ \times 7.6^\circ$. The amplitude spectrum of each individual face was replaced with the average amplitude spectrum of the ten faces prior to the experiment. For more details on stimulus generation, see Gold *et al.* (1999).

The orientation information available to observers during the identification task was manipulated by filtering the stimulus matrices in the spatial frequency domain. In particular, spatial frequency components from a target face were incorporated with components from a standard face that was created by averaging the 10 possible target faces (Figure 3.1). On each trial, an ideal band-pass orientation filter with a full bandwidth (w) ranging from 10° to 180° was used to isolate the frequency components in the average face within $\pm w/2^\circ$ of 0° (horizontal) or 90° (vertical). These components in the average face were then replaced by the corresponding components from a randomly-selected target face. The amplitudes of the frequency components from the target face were scaled so that the total power of the hybrid image was constant across conditions. Note that only frequency components within the passband of the orientation filter contained information about the target face, and that

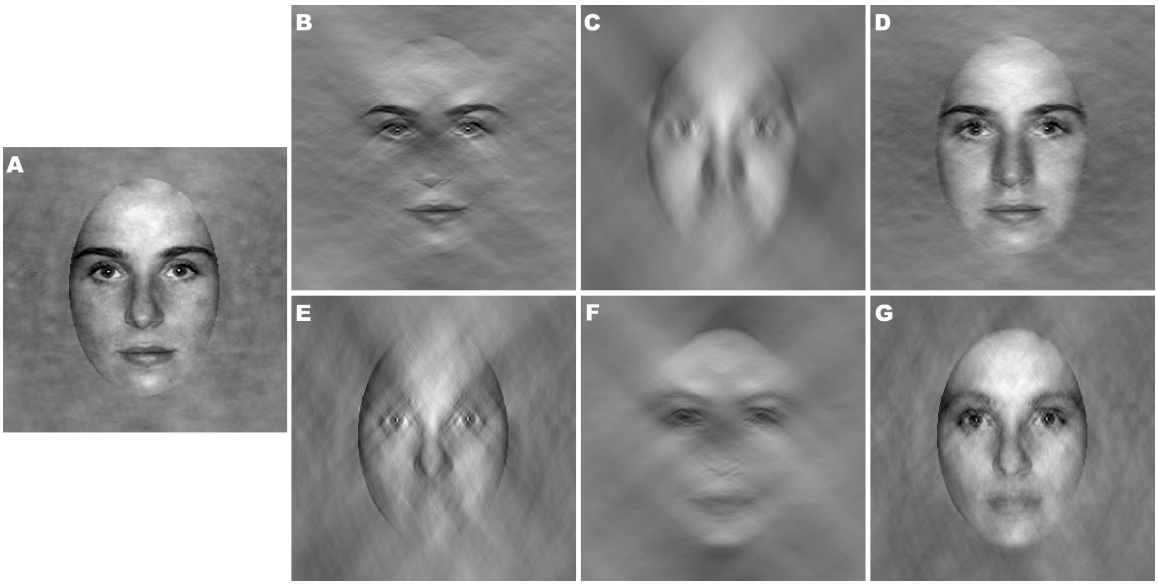


Figure 3.1: **Demonstration of stimulus generation. (A)** An unmanipulated target face. **(B,E)** The target face filtered to retain horizontal (B) or vertical (E) structure ($bandwidth = 90\,^\circ$). **(C,F)** The pixel-wise average of the 10 possible faces filtered to retain vertical (C) or horizontal (F) structure ($bandwidth = 90\,^\circ$). **(D,G)** Final stimuli presented in the experiment, drawn from the horizontal (D) or vertical (G) $90\,^\circ$ bandwidth conditions. These final stimuli were constructed as the pixel-wise sum of the two images to their left.

a 180° filter yielded a stimulus that was identical to the unfiltered target face. This procedure allowed us to manipulate the band of orientations that conveyed stimulus information diagnostic for the identification task while ensuring that stimuli in all conditions looked equally like faces, containing stimulus-typical power in all orientation bands. The final face stimuli, which had an RMS contrast of 0.5, were embedded in a low-contrast white Gaussian noise (RMS contrast = 0.1) prior to presentation. This noise was added to facilitate eventual comparison of data from human observers to that of an ideal observer for which external noise is required.

3.2.3 Design

The experiment consisted of two blocks of trials: face orientation (upright or inverted) varied across blocks, and filter orientation (horizontal or vertical) and bandwidth ($10 - 180°$) varied within blocks. Each block consisted of 360 trials (2 filter orientations $\times$ 18 bandwidths $\times$ 10 repetitions). Prior to each block, observers completed 10 practice trials with unfiltered faces at the appropriate orientation. The session took approximately 45 minutes to complete. The dependent measure was proportion correct on the 10-AFC identification task.

3.2.4 Procedure

Observers viewed the display binocularly, and a chin/head rest was used to stabilize viewing position. Each trial consisted of a fixation cross at the centre of the screen for 500 ms, a 250 ms blank screen, a 250 ms stimulus presentation, another 250 ms blank screen, and finally a response window containing the 10 unfiltered faces that remained available until a response was made. The orientation and size of the faces

on the response screen was identical to that of the target. Observers selected their response using a mouse click, and feedback was provided on every trial using 600 Hz and 200 Hz tones to indicate correct and incorrect responses, respectively.

3.3 Results

3.3.1 Human Observers

All statistical analyses were conducted with R (R Core Team, 2015). Figure 3.2 plots proportion correct in the 10-AFC identification task as a function of filter bandwidth, separately for each filter and face orientation. Lines represent psychometric functions fit to these data using generalized linear models with a probit (i.e., inverse cumulative normal) link function and an additional free parameter, λ, corresponding to the upper asymptote. Note that the largest bandwidth at which the horizontal and vertical filters isolate non-overlapping sets of frequency components is 90 °. As bandwidth increases beyond this point, the frequency components passed by the horizontal and vertical filters become increasingly similar and so, not surprisingly, proportion correct in the two conditions converges to a common point. Visual inspection of Figure 3.2 indicates that performance improved more rapidly as a function of bandwidth for horizontal than vertical filters, and that this difference was larger for upright relative to inverted faces. The maximum difference between accuracy in the horizontal and vertical filter conditions was observed with filter bandwidths of 90 ° with differences of 0.37 and 0.29 for upright and inverted faces, respectively.

For the subsequent analyses, we fitted psychometric functions separately for each

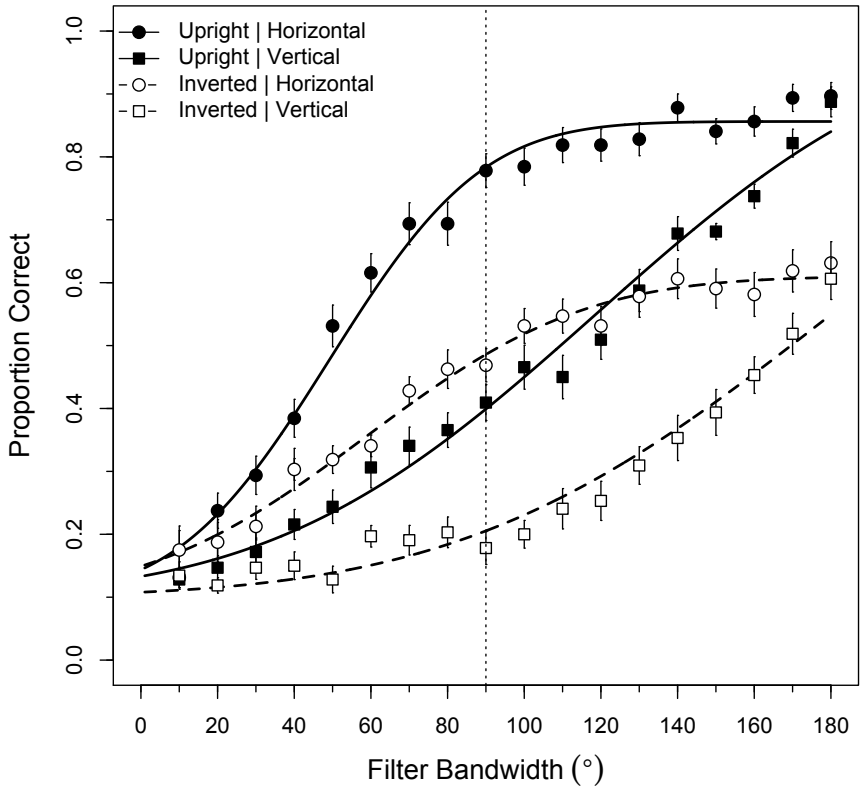


Figure 3.2: Proportion correct in the 10-AFC identification task plotted as a function of filter bandwidth for each face and filter orientation. A full bandwidth of 90 ° is indicated by the vertical dotted line. Lines indicate best-fitting psychometric functions fit to the data using generalized linear models with a probit link function with an additional free parameter λ corresponding to the upper asymptote. When the filter bandwidth was 180 °, the stimuli in the horizontal and vertical filter conditions were identical and corresponded to original (i.e. unfiltered) faces. Error bars represent ± 1 SEM.

observer and extracted measures from these fits. First, we computed the mean proportion correct for each filter and face orientation from 0° to 90°. These values, plotted in Figure 3.3, quantify differences in performance across the full range of orthogonal bandwidths. A 2 (face orientation) × 2 (filter orientation) ANOVA on these values revealed significant main effects of face orientation [$F(1, 31) = 112.97, p < 0.0001$], filter orientation [$F(1, 31) = 160.42, p < 0.0001$], and a significant interaction [$F(1, 31) = 11.25, p = 0.002$]. Together, these results reflect higher performance for upright faces than inverted (main effect of face orientation), higher performance for horizontal structure than vertical (main effect of filter orientation), and more horizontal selectivity for upright faces than inverted (significant interaction), consistent with previous studies (Goffaux and Dakin, 2010; Pachai *et al.*, 2013).

Next, to explore the bandwidth at which horizontal selectivity emerges, we computed separately for each face orientation and filter bandwidth a measure of horizontal selectivity ($PC_{horizontal} - PC_{vertical}$). The resulting values are plotted in Figure 3.4. Visual inspection of Figure 3.4 indicates that horizontal selectivity increases as a function of bandwidth up to 90°, after which it begins to decrease for both face orientations as the two orientation filters grow increasingly similar with bandwidths beyond this point. Note that horizontal selectivity differs as a function of face orientation primarily in the 50-90° bandwidth range. We quantified these results with a 2 (face orientation) × 18 (filter bandwidth) repeated-measures ANOVA conducted on horizontal selectivity. This ANOVA revealed a nearly significant main effect of face orientation [$F(1, 31) = 3.95, p = 0.0558$], a significant main effect of filter bandwidth [$F(17, 527) = 21.39, p < 0.0001$], and a significant face orientation × filter bandwidth interaction [$F(17, 527) = 1.68, p = 0.0435$], confirming the above observations. To our

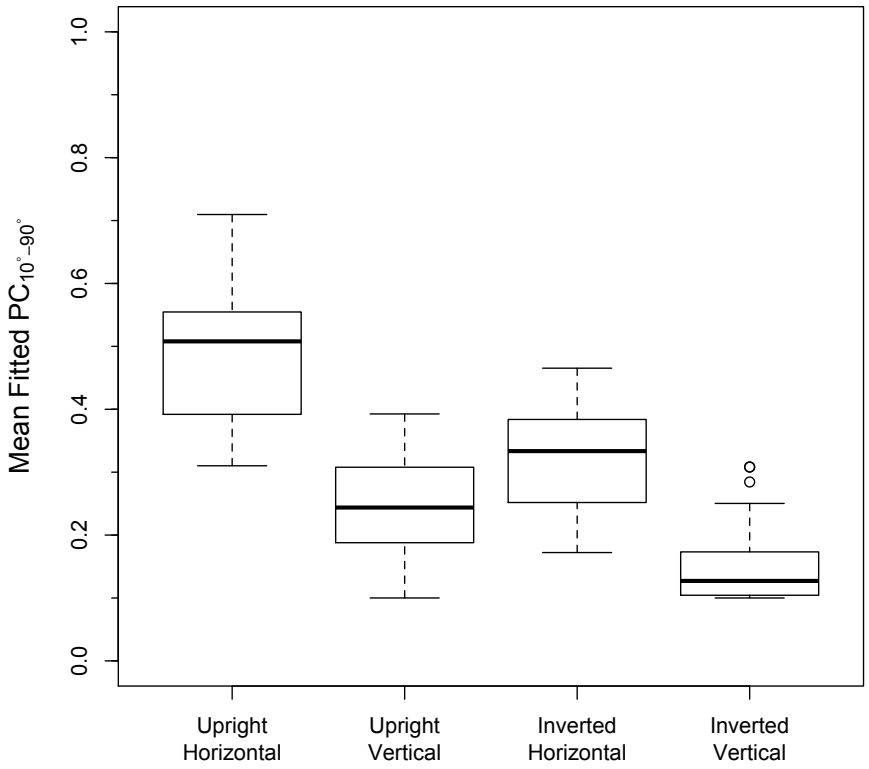


Figure 3.3: Boxplots representing mean proportion correct from 10° to 90° averaged from best-fitting psychometric functions fit separately to each observer's data for each face and filter orientation. The horizontal line in each box represents the median, whereas the upper and lower edges represent the 75th and 25th percentiles, respectively.

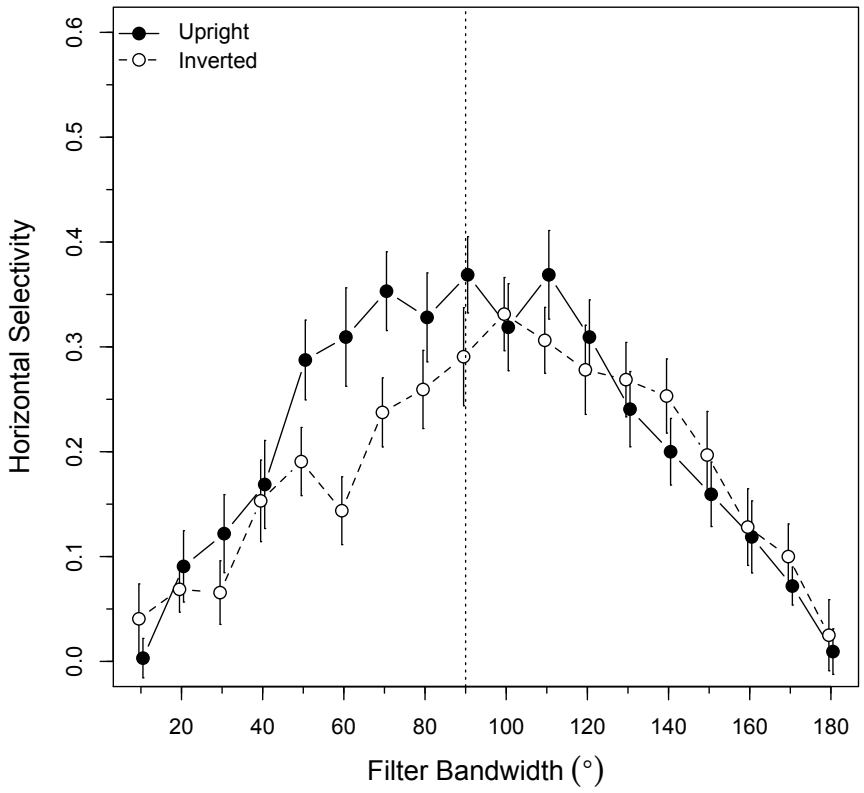


Figure 3.4: Horizontal selectivity, defined as $PC_{horizontal} - PC_{vertical}$, plotted as a function of filter bandwidth separately for each face orientation. A full bandwidth of 90 ° is indicated by the vertical dotted line. Error bars represent ±1 SEM.

knowledge, these results are the first to quantify horizontal selectivity in this way.

3.3.2 Simulated Observers

Ideal Observer

Our human observers demonstrate a clear reliance on information conveyed by horizontal structure to identify upright, and to a lesser extent, inverted faces. However, we know there is more diagnostic information available for identification in the horizontal band (Dakin and Watt, 2009; Pachai *et al.*, 2013). To quantify the extent to which our results reflect the horizontal structure inherent to our stimuli, we simulated the performance of an ideal observer that utilized optimally the available information. For an object identification task such as ours, such an ideal observer correlates the noisy, filtered stimulus with perfect templates of each identity and selects the template yielding the highest correlation (Tjan *et al.*, 1995). This observer isolates precisely the diagnostic orientation band on a given trial and uses optimally the information therein.

The dependent measure of our simulation was proportion correct on the 10-AFC identification task. However, using the same signal-to-noise ratio presented to human observers would result in ceiling performance in all conditions. Therefore, we simulated performance with each filter orientation and bandwidth at seven stimulus RMS contrasts ranging from 0.001 to 0.0025. These contrast values were chosen in pilot simulations to produce neither ceiling nor floor performance in any filter condition. As in the human experiment, Gaussian white noise (RMS contrast = 0.1) was added to the stimulus on every trial. We transformed the resulting proportion correct values into d' using the procedures outlined in Macmillan and Creelman (2004), where

each d' value was based on 5000 simulated trials. Then, separately for each filter orientation and bandwidth, we calculated the best fitting least-squares regression line relating d' to log-transformed RMS contrast (all $R^2 \geq 0.99$) and extrapolated to the RMS contrast of 0.5 presented to human observers.

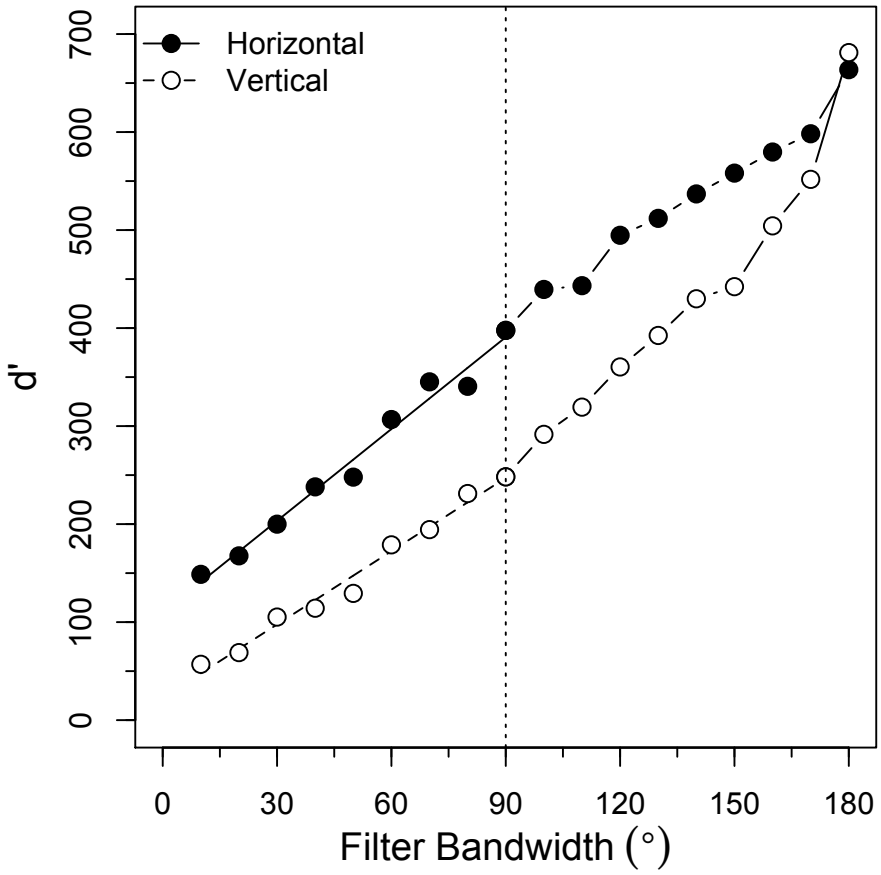


Figure 3.5: d' as a function of filter bandwidth for a simulated observer using a template matching decision rule. Each point is based on 5000 simulated trials. Lines represent the best fitting least-squares regression from 10 to 90 degrees bandwidth. A full bandwidth of 90 ° is indicated by the vertical dotted line.

Extrapolated d' values for our simulated observer are plotted as a function of

Orientation	R^2	Intercept	Slope
Horizontal	0.98	109.67	3.12
Vertical	0.98	22.87	2.49

Table 3.1: Parameters of the best fitting least-squares regression line fit to d' as a function of filter bandwidth from 10 to 90 degrees for a template-matching simulated observer (see text for details)

bandwidth in Figure 3.5. From these data we calculated, separately for each filter orientation, the best fitting least-squares regression line relating d-prime to filter bandwidth from 10 to 90 degrees bandwidth. The resulting parameters are shown in Table 3.1. This analysis revealed a comparatively small difference in slopes between horizontal and vertical filters and a large difference between intercepts. The intercepts indicate that the differential value of information in the horizontal and vertical bands lies in the narrowest bandwidth range (< 10 degrees), while the slopes indicate the informational gain of increasing bandwidth is roughly equal for the two filters, although slightly higher in the horizontal band. To demonstrate this effect more directly, we computed horizontal selectivity for the ideal observer as the difference in d' between horizontal and vertical filters, similarly to the human data plotted in Figure 3.4. The resulting values, plotted in figure 3.6, reveal relatively high and constant horizontal tuning from 10° to 90° reflecting the horizontal advantage conveyed at a bandwidth narrower than 10°.

Absolute Efficiency

We next examined the relationship between the performance of our human observers and the ideal observer. To this end, we transformed individual observer's data to d'

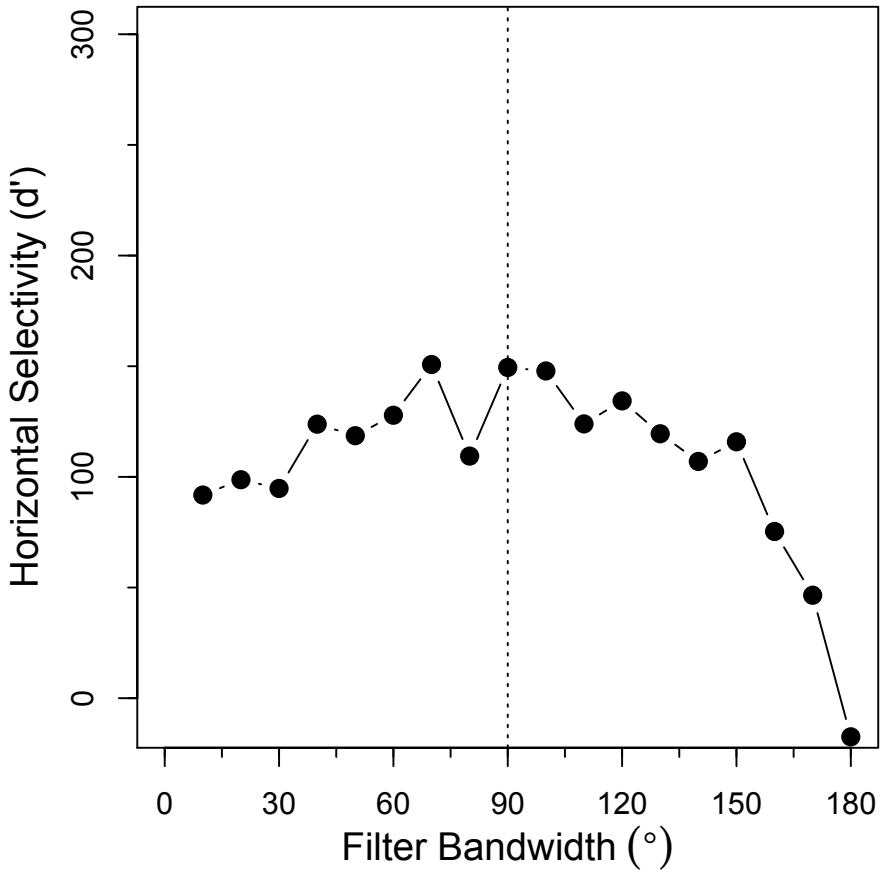


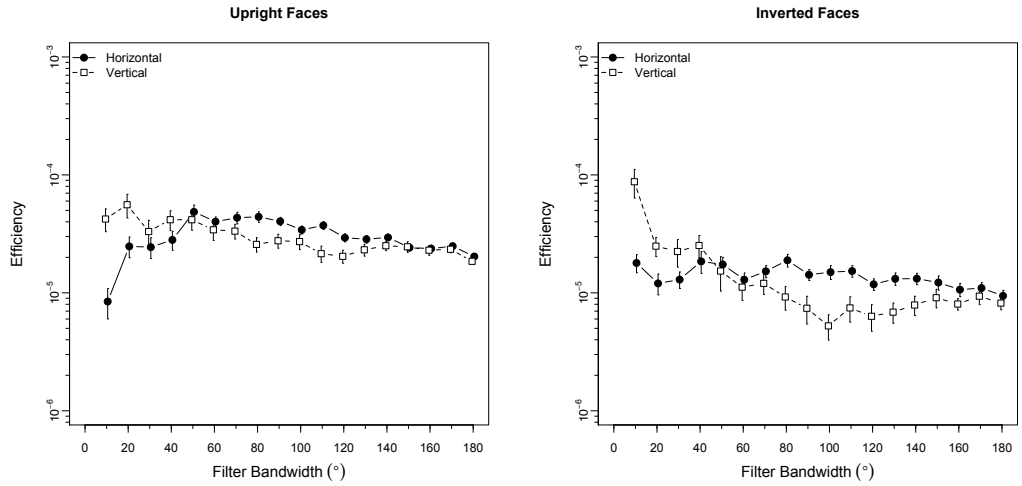
Figure 3.6: Horizontal Selectivity for the ideal observer, calculated as $d'_{horizontal} - d'_{vertial}$, plotted as a function of filter bandwidth. A full bandwidth of 90 ° is indicated by the vertical dotted line.

values, converting accuracy of 0 and 1 to 0.05 and 0.95 respectively, based on the procedures described in Macmillan and Creelman (2004). We then calculated, separately for each observer, efficiency (η) which is equal to $(d'_{human}/d'_{ideal})^2$ (Tanner and Birdsall, 1958). On this measure, a value of 1 indicates that observers are performing optimally, while lower values indicate less optimal performance.

Efficiency is plotted as a function of bandwidth separately for each face and filter orientation in Figure 3.7. We first note the low overall values of efficiency, common in face identification tasks (Gaspar *et al.*, 2008a; Gold *et al.*, 1999; Pachai *et al.*, 2013). To quantify these results further, we conducted a 2 (face orientation) $\times$ 2 (filter orientation) $\times$ 9 (filter bandwidth, 10°-90°) repeated-measures ANOVA. This analysis revealed significant main effects of face orientation $[F(1, 31) = 41.69, p < 0.0001]$ and filter bandwidth $[F(8, 248) = 3.11, p = 0.0023]$. However, the main effect of filter orientation was not significant $[F(1, 31) = 3.94, p = 0.0559]$, nor was the face orientation $\times$ filter orientation interaction $[F(1, 31) = 0.79, p = 0.3811]$. The filter orientation $\times$ filter bandwidth interaction was significant $[F(8, 248) = 11.91, p < 0.0001]$, as was the face orientation $\times$ filter bandwidth interaction $[F(8, 248) = 7.28, p < 0.0001]$. Finally, the face orientation $\times$ filter orientation $\times$ filter bandwidth interaction was not significant $[F(8, 248) = 1.36, p = 0.2156]$. Note that these analyses were conducted on bandwidths from 10° to 90° because beyond 90° the filters begin to overlap, but an identical ANOVA conducted from 10°-180° revealed qualitatively identical results.

Fixed-Bandwidth Observer

Finally, we simulated the performance of a suboptimal observer that was unable to isolate optimally the diagnostic information on a given trial. Recall that horizontal



(a) (b)

Figure 3.7: Efficiency relative to an ideal observer is plotted as a function of filter bandwidth for upright faces (left) and inverted faces (right). See text for details. Error bars represent ±1 SEM.

selectivity in human observers increased progressively with bandwidth until approximately 60° - 90° (Figure 3.4). For this reason, we simulated an observer that utilized optimally the horizontal structure conveyed in one of three bandwidths (30°, 60°, or 90°) while utilizing suboptimally all other orientation components (30% of the optimal weight). This observer used the same decision rule described in Section 3.3.2. As with the ideal observer in Section 3.3.2, we simulated the performance of these observers at seven stimulus RMS contrasts ranging from 0.001 to 0.0025 before extrapolating to the RMS contrast shown to human observers. Visual inspection of the resulting values, plotted in Figure 3.8 reveal a similar pattern of results to human observers. In particular, the suboptimal use of vertical structure imposed on these observers results in a progressive increase in horizontal selectivity as filter bandwidth approaches the bandwidth preferred by each observer. This pattern of horizontal selectivity is

plotted directly in Figure 3.9. Although it is difficult to determine precisely the fixed bandwidth that best approximates the performance of human observers, it is clear that relatively suboptimal use of vertical structure characterizes human performance more accurately than an optimal observer (i.e., Section 3.3.2).

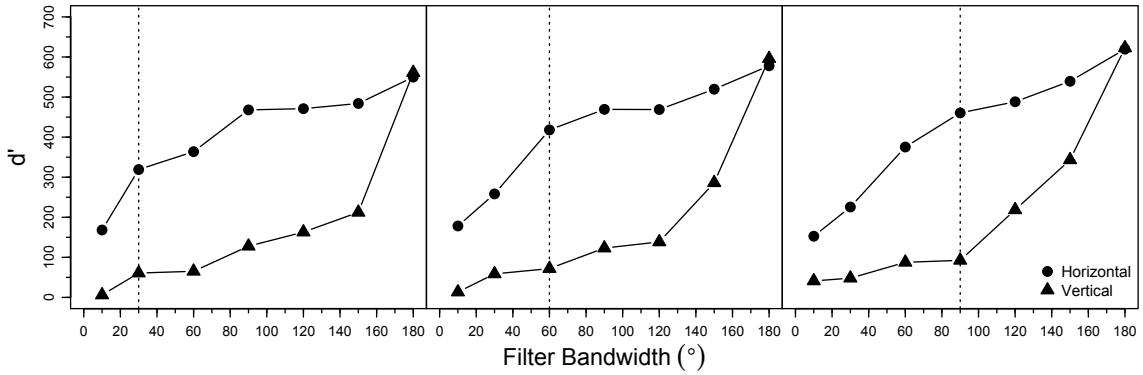


Figure 3.8: d' as a function of filter bandwidth for three simulated observers using a template matching decision rule. Each observer bases its decision on a fixed bandwidth of horizontal structure (30°, 60°, or 90°, indicated by the dotted line) while utilizing all other orientation components suboptimally (30% of the optimal weight). Each point is based on 5000 simulated trials.

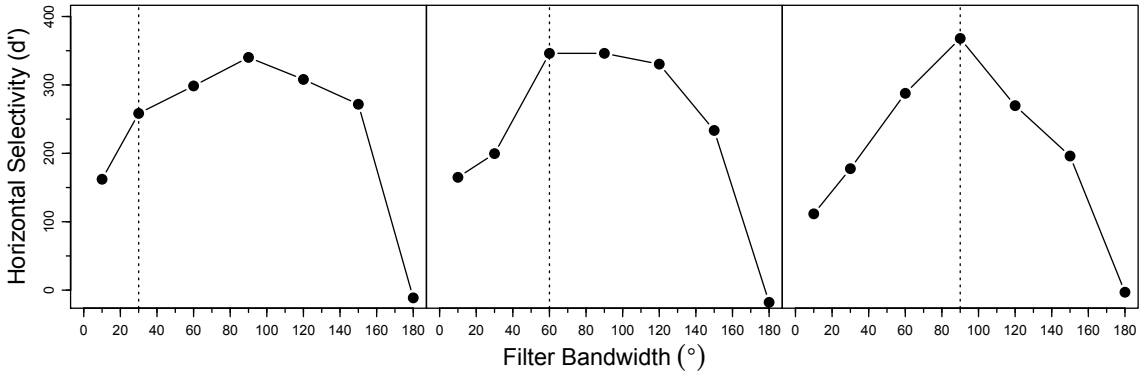


Figure 3.9: Horizontal Selectivity, , calculated as $d'_{horizontal} - d'_{vertial}$, for each of the simulated observers shown in Figure 3.8. Dotted lines indicate the bandwidth, centred on the horizontal band, used optimally by each observer.

3.4 Discussion

The results of this experiment are consistent with recent reports that preferential use of horizontal structure supports accurate identification of upright faces, and that this preference is reduced by picture plane inversion (Dakin and Watt, 2009; Goffaux and Dakin, 2010; Goffaux *et al.*, 2011; Pachai *et al.*, 2013). Further, we have quantified the bandwidth at which the horizontal bias inherent to face stimuli is conveyed, and shown that the horizontal selectivity demonstrated by human observers is driven by efficient use of this information. In Figure 3.5, we show that an ideal observer performs better when provided horizontal structure than vertical at 10° bandwidth, and that the magnitude of this difference is relatively constant as bandwidth increases. This result suggests that the orientation structure diagnostic for face identification is especially conveyed along the horizontal meridian with a bandwidth less than 10°. At such narrow bandwidths, human observers perform near chance for all filter conditions (see Figure 3.2). Consequently, absolute efficiency is *lower* for horizontal structure than vertical in this range, presumably because real observers integrate information over a larger bandwidth range. However, horizontal selectivity increases rapidly with bandwidth (see Figure 3.4) until the information contained in horizontal and vertical bandwidths converges to a common point for 180°. Together, these results suggest that human observers integrate information across a wider orientation range than the narrow band in which the most diagnostic horizontal structure is conveyed. Indeed, when we simulated observers with similar constraints (Section 3.3.2), the pattern of results matched human performance more closely. These results are consistent with previous psychophysical estimates of minimum orientation channel bandwidths, ranging from $11 - 30°$ (see Govenlock *et al.*, 2009 for a review). However, the precise

manner in which early orientation signals are combined (e.g., Beaudot and Mullen, 2006) for face identification remains a question for future research.

Although the diagnostic structure of horizontal orientations remains constant when faces are inverted, we observed a significant face inversion effect, with lower overall identification for inverted faces than for upright faces. As seen in Figure 3.2, the magnitude of the face inversion effect increases with orientation bandwidth until about 90°, concomitant with the zone of reduced horizontal selectivity for inverted faces compared to upright faces (Figure 3.4). Efficiency for both horizontal and vertical structure also was reduced for inverted faces, except at the very smallest orientation bandwidth (Figure 3.7). Taken as a whole, these results suggest a limited extent to which the visual system can adjust the orientation tuning characteristics of visual channels to optimize performance based on the availability of physical information in the stimulus (Taylor *et al.*, 2014). Given that perceptual learning has been shown to alter the magnitude of the face inversion effect (Hussain *et al.*, 2009; Laguesse *et al.*, 2012), future research should examine whether training increases orientation channel flexibility, specifically targeting the selective processing of horizontal structure, or whether overall efficiency for all orientations simply increases with training.

Selective extraction of horizontal structure has been shown to affect other robust phenomena associated with face perception. For example, Goffaux and Dakin (2010) demonstrated that when horizontal structure is rendered unavailable, identity aftereffects and holistic processing are significantly reduced. Face perception also is associated with neural signals such as the N170, a negative-going peak in the ERP that begins to emerge 130 ms after presentation of a face (Rousselet *et al.*, 2008).

Jacques *et al.* (2014) reported delayed N170 peaks and an attenuated N170 inversion effect when horizontal structure was phase-randomized, which they attribute to impaired recruitment of face-specific neural representations when horizontal information is rendered unavailable. Further, Huynh and Balas (2014) demonstrated that horizontal structure is required for accurate discrimination of happy and sad facial expressions. Interestingly, in the same study, when the faces were rotated 90 ° such that the horizontal structure became vertical in image-centred coordinates, observers remained tuned to the horizontal information in object-centred coordinates, suggesting some flexibility in the orientation channels used during face perception. Future research should determine more precisely the limits of that flexibility, and the behavioural timeframe over which selectivity for horizontal structure emerges.

Overall, the current results reiterate the important role of horizontal structure for face identification. We demonstrate that a narrow band of orientation components along the horizontal meridian is maximally diagnostic for face identification, but that human observers utilize a much wider range of information when performing such tasks.

Bibliography

Beaudot, W. H. A. and Mullen, K. T. (2006). Orientation discrimination in human vision: psychophysics and modeling. *Vision Research*, **46**(1-2), 26–46.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**(4), 433–6.

Dakin, S. C. and Watt, R. J. (2009). Biological bar codes in human faces. *Journal of Vision*, **9**(4), 1–10.

Diamond, R. and Carey, S. (1986). Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology: General*, **115**(2), 107–17.

Farah, M. J., Wilson, K. D., Drain, H. M., and Tanaka, J. W. (1998). What is "special" about face perception? *Psychological review*, **105**(3), 482–98.

Gaspar, C. M., Sekuler, A. B., and Bennett, P. J. (2008a). Spatial frequency tuning of upright and inverted face identification. *Vision Research*, **48**(28), 2817–26.

Gaspar, C. M., Bennett, P. J., and Sekuler, A. B. (2008b). The effects of face inversion and contrast-reversal on efficiency and internal noise. *Vision Research*, **48**(8), 1084–95.

Gauthier, I. and Tarr, M. J. (1997). Becoming a "Greeble" Expert: Exploring Mechanisms for face Recognition. *Vision Research*, **37**(12), 1673–1682.

Goffaux, V. and Dakin, S. C. (2010). Horizontal information drives the behavioral signatures of face processing. *Frontiers in Perception*, **1**, 1–14.

Goffaux, V., van Zon, J., and Schiltz, C. (2011). The horizontal tuning of face perception relies on the processing of intermediate and high spatial frequencies. *Journal of Vision*, **11**, 1–9.

Gold, J. M., Bennett, P. J., and Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research*, **39**(21), 3537–60.

Gosselin, F. and Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision research*, **41**(17), 2261–71.

Govenlock, S. W., Taylor, C. P., Sekuler, A. B., and Bennett, P. J. (2009). The effect of aging on the orientational selectivity of the human visual system. *Vision research*, **49**(1), 164–72.

Heisz, J. J. and Shore, D. I. (2008). More efficient scanning for familiar faces. *Journal of Vision*, **8**(1), 1–10.

Hussain, Z., Sekuler, A. B., and Bennett, P. J. (2009). Perceptual learning modifies inversion effects for faces and textures. *Vision research*, **49**(18), 2273–84.

Huynh, C. M. and Balas, B. (2014). Emotion recognition (sometimes) depends on horizontal orientations. *Attention, perception & psychophysics*.

Jacques, C., Schiltz, C., and Goffaux, V. (2014). Face perception is tuned to horizontal orientation in the N170 time window. *Journal of Vision*, **14**, 1–18.

Laguesse, R., Dormal, G., Biervoye, A., Kuefner, D., and Rossion, B. (2012). Extensive visual training in adulthood significantly reduces the face inversion effect. *Journal of Vision*, **12**, 1–13.

LeGrand, R., Mondloch, C. J., Maurer, D., and Brent, H. P. (2004). Impairment in Holistic Face Impairment Early Processing Following Visual Deprivation. *Psychological Science*, **15**(11), 762–768.

Macmillan, N. A. and Creelman, C. D. (2004). *Detection Theory: A User's Guide*. Psychology Press, 2nd edition.

Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research*, **39**(23), 3824–3833.

Obermeyer, S., Kolling, T., Schaich, A., and Knopf, M. (2012). Differences between Old and Young Adults' Ability to Recognize Human Faces Underlie Processing of Horizontal Information. *Frontiers in aging neuroscience*, **4**(April), 3.

Pachai, M. V., Sekuler, A. B., and Bennett, P. J. (2013). Sensitivity to Information Conveyed by Horizontal Contours is Correlated with Face Identification Accuracy. *Frontiers in Psychology*, **4**(February), 1–9.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, **10**(4), 437–442.

Peterson, M. F. and Eckstein, M. P. (2012). Looking just below the eyes is optimal

across face recognition tasks. *Proceedings of the National Academy of Sciences*, **109**(48).

R Core Team (2015). R: A Language and Environment for Statistical Computing.

Ringach, D. L., Shapley, R. M., and Hawken, M. J. (2002). Orientation selectivity in macaque V1: diversity and laminar dependence. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **22**(13), 5639–51.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta psychologica*, **128**(2), 274–89.

Rousselet, G. A., Pernet, C. R., Bennett, P. J., and Sekuler, A. B. (2008). Parametric study of EEG sensitivity to phase noise during face processing. *BMC Neuroscience*, **9**, 98.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., and Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, **14**(5), 391–6.

Tanaka, J. (2001). The entry point of face recognition: evidence for face expertise. *Journal of Experimental Psychology: General*, **130**(3), 534–543.

Tanaka, J. W. and Farah, M. J. (1993). Parts and Wholes in Face Recognition. *Journal of Experimental Psychology*, **46**(2), 225–245.

Tanner, W. P. and Birdsall, T. G. (1958). Definitions of d' and eta as psychophysical measures. *The Journal of the Acoustical Society of America*, **30**(10), 922–928.

Taylor, C. P., Bennett, P. J., and Sekuler, A. B. (2014). Evidence for adjustable bandwidth orientation channels. *Frontiers in Psychology*, **5**, 1–10.

Tjan, B. S., Braje, W. L., Legge, G. E., and Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research*, **35**(21), 3053–69.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, **79**(4), 471–491.

Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J. W., Bub, D. N., and Gosselin, F. (2010). Does face inversion change spatial frequency tuning? *Journal of experimental psychology. Human perception and performance*, **36**(1), 122–35.

Yarbus, A. L. (1967). *Eye Movements During Perception of Complex Objects*. Plenum Press, New York.

Yin, R. K. (1969). Looking at Upside-Down Faces. *Journal of Experimental Psychology*, **81**(1), 141–145.

Yu, D. and Chung, S. T. L. (2011). Critical Orientation for Face Identification in Central Vision Loss. *Optometry and Vision Science*, **88**(6), 724–732.

Chapter 4

Masking of individual facial features reveals the use of horizontal structure in the eyes

4.1 Introduction

The results of many studies suggest that information contained near the eyes is important for identifying human faces. Yarbus (1967) first observed that human fixations to face stimuli are directed to the eyes, nose, and mouth in turn, a result that has been replicated repeatedly (e.g., Arizpe *et al.*, 2012; Blais *et al.*, 2008; Mehoudar *et al.*, 2014; Miellet *et al.*, 2013; Williams and Henderson, 2007). This T-shaped pattern is observed when faces are presented for long durations, but recent evidence suggests that the first 1-2 fixations are sufficient for accurate face identification (Hsiao and Cottrell, 2008). Examination of these initial fixations reveals a bias for the centre

of the face, just below the eyes (Peterson and Eckstein, 2013). This result is consistent with studies using response classification (Sekuler *et al.*, 2004) and bubbles (Gosselin and Schyns, 2001) techniques, which reveal that information from the eye region underlies performance in face identification tasks. Indeed, ideal observer analyses reveal that fixating near the centre of the face is optimal for a foveated observer, because it facilitates extraction of information from both eyes (Peterson and Eckstein, 2012). Further, with increasing familiarity, observers gradually decrease the number and scope of their saccades, focusing on more on the eye region (Althoff and Cohen, 1999; Heisz and Shore, 2008; van Belle *et al.*, 2010). Together, these results suggest that processing of information from the eyes is an important component of face identification.

Preferential sampling associated with identification expertise has also been observed in the Fourier domain. Dakin and Watt (2009) first demonstrated that information conveyed by horizontally-oriented spatial frequency components is particularly diagnostic for face identity, and faces filtered to retain this information are identified more accurately than faces retaining other orientation structure. Following this work, Goffaux and Dakin (2010) demonstrated that horizontal structure must be available to observe face-related phenomena, such as identity aftereffects and holistic processing, and Pachai *et al.* (2013) demonstrated that inter-observer variability in selective processing of horizontal structure is correlated with face identification performance. Sensitivity to horizontal structure also decreases following face inversion (Goffaux and Dakin, 2010; Pachai *et al.*, 2013), and the electrophysiological signatures of face inversion are attenuated when this structure is rendered unavailable (Jacques *et al.*, 2014). Together, these results suggest that horizontal selectivity in the Fourier domain is

another signature of face identification expertise.

In the present study, we explored the relationship between the spatial and Fourier characteristics of human face identification. Specifically, we hypothesized that human observers preferentially process horizontal structure from the eyes during face identification tasks because this is the most diagnostic source of information in the stimulus. To explore this hypothesis, we selectively masked particular facial features using external noise, either unfiltered or band-pass filtered to retain horizontal or vertical structure, and measured the consequent decrement in performance. By measuring the differential effects of these noise conditions in human observers, we quantified the spatial locations at which selective horizontal processing is deployed. Further, by simulating the performance of an ideal observer under identical conditions, we quantified the distribution of orientation information in the stimulus. Finally, by comparing human observers to the ideal observer, we quantified the extent to which human information sampling can be described as optimal.

4.2 Methods

4.2.1 Observers

Twenty observers (10 male) participated in the experiment. Their age range was 18 - 27 (μ = 22). All observers had normal or corrected-to-normal Snellen acuity. Compensation was provided at the rate of $10/hour or as partial course credit. All experimental protocols were approved by the McMaster University Research Ethics Board, and informed consent was collected prior to the experiment.

4.2.2 Apparatus

Experimental sessions were run on an Apple Macintosh G5 computer using MATLAB with the Psychophysics and Video Toolboxes (Brainard, 1997; Pelli, 1997). Stimuli were presented on a calibrated 21" Apple Studio display with a resolution of 1152 $\times$ 870 pixels and a frame rate of 75 Hz. Average luminance was 18.5 cd/m^2. A chin rest stabilized viewing distance at 60 cm, and the monitor was the only light source in the room.

4.2.3 Stimuli

Stimuli were based on digital photographs of five males and five females with no visible piercings, facial fair, or eye glasses. Each face was cropped with a 148 $\times$ 105 pixel oval window that isolated the internal features, and was centred in a 192 $\times$ 192 pixel matrix with a uniform grey background. At the viewing distance of 60 cm, stimulus matrices subtended 6.4° $\times$ 6.4°. Prior to presentation, the amplitude spectrum of each face was adjusted to the average amplitude spectrum for the face set.

During the experiment, face contrast was adjusted using QUEST (Watson and Pelli, 1983) to determine the 50% correct 10-AFC identification threshold. Two independent Gaussian noise fields were added to the stimulus on every trial: a low contrast (RMS = 0.0001) baseline white noise that covered the entire image and a high contrast (RMS = 0.2) masking noise. The baseline noise was included to permit ideal observer analysis. Masking noise was either unfiltered or filtered to isolate horizontal or vertical frequency components (sharp edged, bandwidth = 90°) and passed through Gaussian apertures (full-width at half-maximum = 1.1°) centred on one or

more facial features (left eye, right eye, nose, mouth). See Figure 4.1 for examples of the stimuli.

4.2.4 Design

The experiment comprised two sessions, each including six aperture conditions. In the first session, observers viewed each of the four single-feature aperture conditions, the four-feature aperture combination, and the no-aperture control condition in which the whole stimulus array was masked. In the second session, observers viewed each of the three-feature aperture combinations, the left eye + right eye aperture combination, the nose + mouth aperture combination, and the low-noise baseline condition with no additional mask. On an equal number of trials per aperture condition, the noise was white (i.e., unfiltered), horizontal, or vertical. Filter and aperture conditions were randomly intermixed within sessions. Observers completed 50 trials with each noise/aperture combination, for a total of 1050 trials per session. Prior to each session, observers completed 10 practice trials with high contrast faces (RMS = 0.3) and randomly chosen conditions applicable to the current session.

Each trial began with a fixation cross presented at the centre of the screen for 500 ms. The stimulus was presented for 250 ms, with a 250 ms blank before and after presentation. Finally, observers viewed a selection screen containing high contrast, unfiltered exemplars of each face at the same size as the target. Identity responses were provided using a mouse click. Auditory feedback was provided on every trial. The dependent measure was the 50% correct 10-AFC identification threshold, measured separately for each noise and aperture condition.

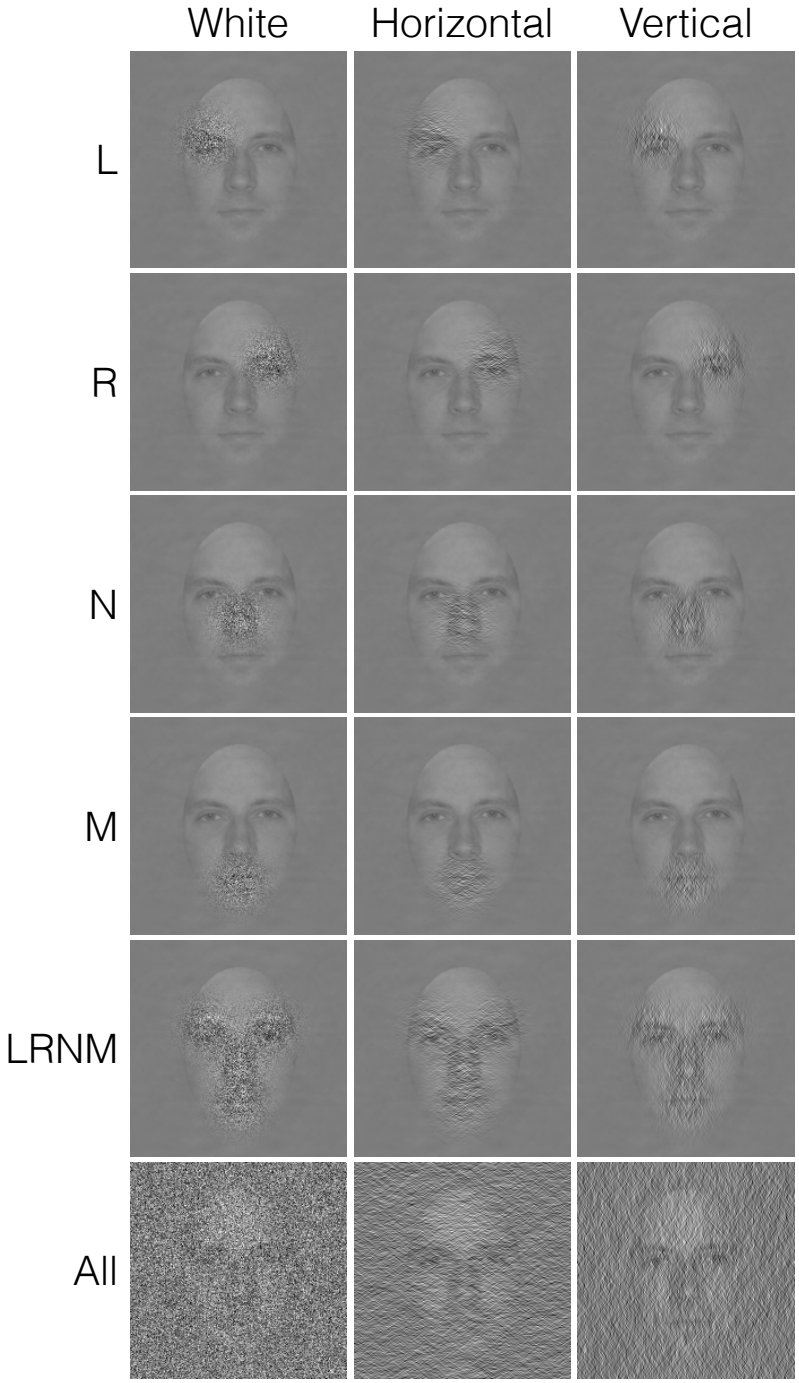


Figure 4.1: Example subset of the experimental conditions. L indicates the left eye, R indicates the right eye, N indicates the nose, and M indicates the mouth. Face and noise contrast increased for visibility.

4.2.5 Ideal Observer Analysis

An ideal observer uses an optimal decision rule to achieve the best possible performance on a task given the available information (Bennett and Banks, 1987; Banks *et al.*, 1991; Tjan *et al.*, 1995). Generally speaking, when a noise mask is applied to the stimulus, ideal performance will be affected to the extent that the masked information was diagnostic for the task at hand. In this way, when we mask individual facial features, the decrement in ideal observer performance reveals the extent to which those features contribute to identity discrimination. Further, when we mask these features with either horizontal or vertical noise, the differential effect on ideal performance reveals the extent to which the diagnostic information at those locations is selectively conveyed by horizontal frequency components.

The ideal observer for object identification tasks that use white noise is generally a template matcher that measures the *a posteriori* probability of each identity given the noisy target. More specifically, if R is the noisy stimulus, σ^2 is the variance of the noise, T_i is the noiseless template of the ith identity, and $P(T_i)$ is the *a priori* probability of T_i, then the ideal observer maximizes the function

$$exp\left(-\frac{1}{2\sigma^2}\left\|R-T_i\right\|^2\right)P(T_i) \tag{4.1}$$

by minimizing $\|R-T_i\|^2$, which is the Euclidian distance between stimulus and template. The ideal observer has also perfect information about the stimulus manipulations on a trial-by-trial basis and adjusts its decision rule accordingly. Because our experiment used orientation-filtered noise, we adjusted the ideal templates by computing the product, in the Fourier domain, of the original templates and a pre-whitening filter that reduced the weight given to the masked orientation bands (Conrey and

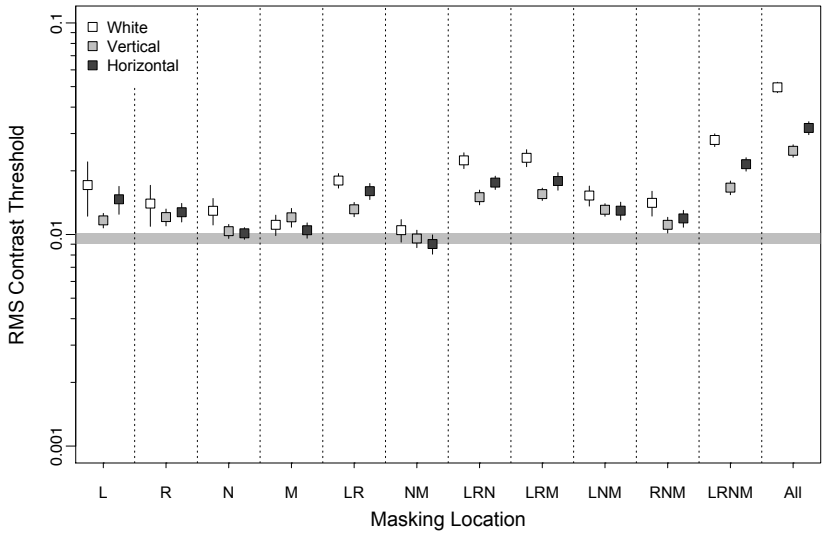
Gold, 2009). Further, because our noise was passed through gaussian apertures that selectively masked particular spatial locations, we adjusted the ideal templates by computing the product of the original templates and inverse apertures that reduced the weight given to the masked locations. These adjusted templates were used to maximize Equation 4.1.

4.3 Results

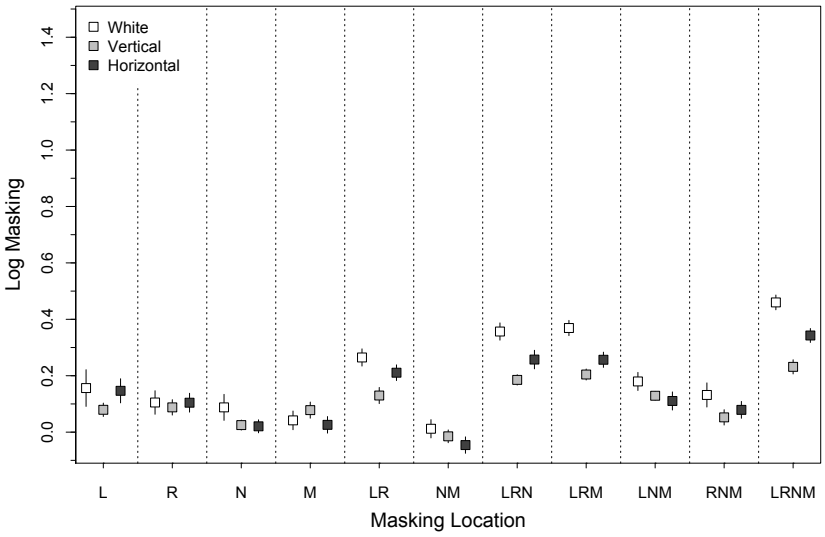
We began by analyzing the thresholds obtained in the low-noise baseline condition. The QUEST threshold estimation procedure generates a posterior probability distribution for each threshold, with the mean representing the threshold estimate and standard deviation representing the estimate's reliability. We estimated three such thresholds per observer in the baseline condition and used their mean as our final measure. This measure was considered stable if all three thresholds fell within $\pm$ 3 standard deviations of their mean. Using this criterion, five observers were rejected for producing unstable baseline threshold measures, leaving a final n of 15 observers for all subsequent analyses.

Figure 4.2a plots the mean thresholds measured for human observers. Recall that for every condition but the baseline, particular features were masked with white noise, horizontally-filtered noise, or vertically-filtered noise. In this design, the lower bound on thresholds is revealed by the baseline condition. For each observer, we used this bound to transform thresholds into log masking using the formula:

$$\text{masking} = \log_{10}\left(\frac{c_{masked}}{c_{baseline}}\right) \tag{4.2}$$



(a)



(b)

Figure 4.2: **(A)** Mean RMS contrast thresholds plotted as a function of masking condition. Masking conditions comprise combinations of the left eye (L), right eye (R), nose (N) and mouth (M) as well as a whole-stimulus mask (All). The horizontal bar represents the mean baseline threshold ± 1 SEM. **(B)** Log masking, where a value of 0 represents the mean threshold obtained in the low-noise baseline condition. In both figures, error bars represent ± 1 SEM.

where c represents the RMS contrast thresholds. These data are plotted in Figure 4.2b. On this measure, the masking observed in the white noise conditions quantifies observers' reliance on the masked features, whereas the difference in masking between horizontal and vertical noise quantifies observers' selective processing of horizontal structure from these features. White noise masking and horizontal-vertical differences in the various conditions were evaluated using t tests (see Table 4.1). Masking was greatest in conditions in which the mask covered both eyes, indicating that observers relied on information around the eyes to perform the task. Also, horizontal selectivity only emerged when both eyes were masked, indicating that horizontal structure around the eyes is most important for identification. We further quantified these results using a 2 (filter, horizontal or vertical) $\times$ 12 (aperture, one per condition) repeated-measures ANOVA, which revealed significant main effects of filter [$F(1, 14) = 27.57, p = 0.0001$], demonstrating horizontal selectivity, aperture [$F(11, 154) = 64.27, p < 0.0001$], demonstrating spatial selectivity, and a significant filter $\times$ aperture interaction [$F(11, 154) = 4.49, p < 0.0001$], demonstrating that horizontal selectivity differed across the face. To quantify whether the horizontal selectivity observed for whole-face masking was driven by extraction of horizontal structure from the eyes, we submitted all aperture conditions in which both eyes were masked (LR, LRN, LRM, LRNM, All) to a 2 (filter orientation) $\times$ 5 (aperture condition) repeated-measures ANOVA. This analysis revealed significant main effects of filter [$F(1, 14) = 28.78, p < 0.0001$] and aperture [$F(11, 154) = 11.41, p < 0.0001$], but no filter $\times$ aperture interaction [$F(11, 154) = 1.19, p = 0.32$]. This non-significant interaction suggests that preferential processing of horizontal structure from the eyes underlied the horizontal selectivity observed when additional features are masked, as

selectivity did not vary significantly across these conditions.

	white noise masking				horizontal selectivity			
	μ	t	df	p	μ	t	df	p
L	0.156	1.46	14	0.083	0.067	1.47	14	0.081
R	0.105	1.72	14	0.053	0.017	0.57	14	0.289
N	0.088	1.94	14	0.037	-0.004	-0.16	14	0.564
M	0.042	1.33	14	0.102	-0.052	-1.17	14	0.869
LR	0.265	6.05	14	2e-05	0.081	3.22	14	0.003
NM	0.012	0.80	14	0.218	-0.031	-1.36	14	0.902
LRN	0.357	8.66	14	3e-07	0.072	2.43	14	0.015
LRM	0.369	9.58	14	8e-08	0.052	2.27	14	0.019
LNM	0.180	4.33	14	3e-04	-0.019	-0.52	14	0.695
RNM	0.132	3.26	14	0.003	0.027	0.86	14	0.203
LRNM	0.460	13.1	14	2e-09	0.111	3.72	14	0.001
All	0.713	30.5	14	2e-14	0.107	4.41	14	3e-04

Table 4.1: **Left:** one-tailed t-tests evaluating the hypothesis that normalized masking in the white noise condition was significantly greater than 0. **Right:** paired one-tailed t-tests evaluating the hypothesis that masking in the horizontal noise condition was significantly greater than masking in the vertical noise condition, a pattern we term horizontal selectivity. Masked features are indicated across rows, where L indicates the left eye, R indicates the right eye, N indicates the nose, and M indicates the mouth.

We next examined the thresholds measured for the ideal observer. As with the human observers, we represent these data as both RMS contrast thresholds and normalized masking in Figure 4.3. For the ideal observer, the masking observed in the white noise conditions quantifies the information diagnostic for identity conveyed at each feature location, whereas the difference between the horizontal and vertical masking conditions quantifies the extent to which this information is selectively conveyed by horizontal frequency components. Examination of Figure 4.3 reveals that the eyes

convey more diagnostic information than the noise and mouth and that the diagnostic information in the eyes and mouth, but not the nose, is selectively conveyed by horizontal frequency components.

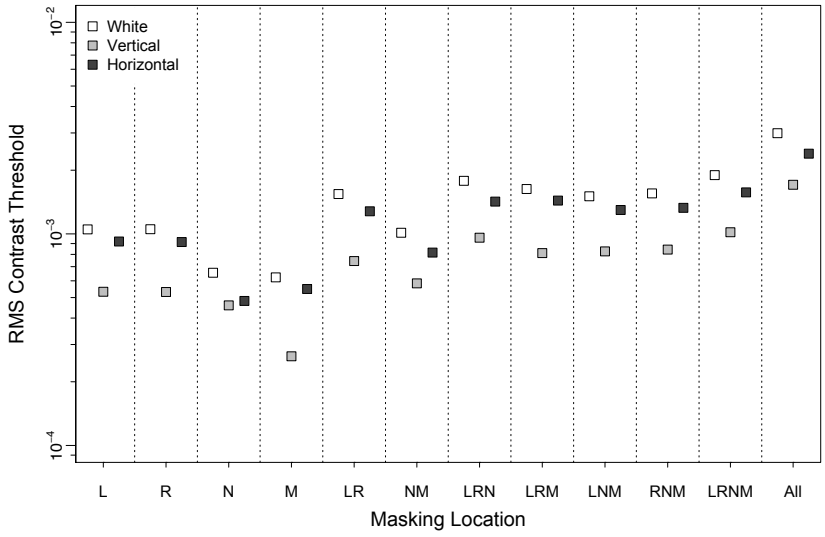
Having quantified the information content of the stimulus using an ideal observer, we next explored the extent to which the pattern of human results can be described as optimal. One measure of human performance relative to an ideal observer is calculated using the formula:

$$\eta = \frac{E_{ideal}}{E_{human}} \tag{4.3}$$

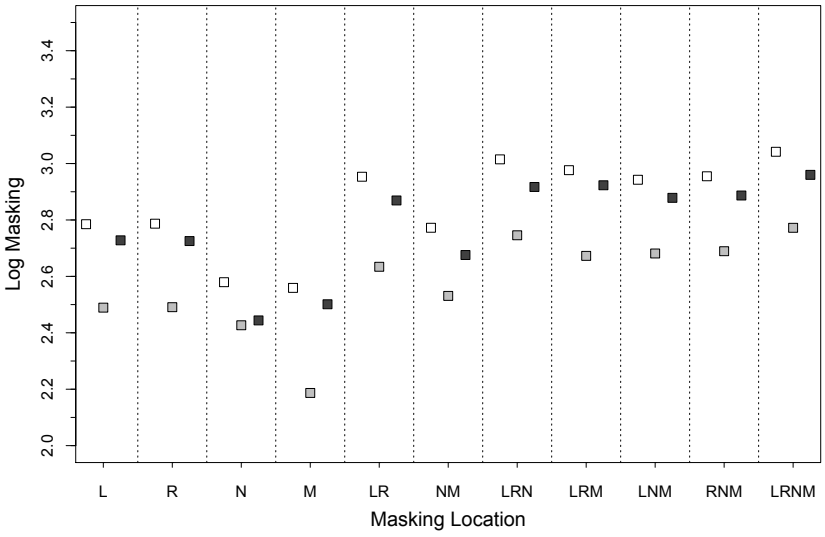
where η is termed efficiency (Tanner and Birdsall, 1958) and E is termed contrast energy, defined as:

$$E = c_{RMS}^2 na \tag{4.4}$$

where c_{RMS} is the contrast threshold, n is the number of pixels in the image and a is the area of a single pixel in deg^2. Usually, higher values of η indicate more efficient, or optimal performance. However, in a masking paradigm such as the one used in the current experiment, higher values of η can indicate *less* optimal use of the masked information in conditions where the ideal observer is more affected by the mask than human observers, which results in a reduction of the difference between ideal and human thresholds. Consider an illustrative case where a mask increases the ideal threshold so extensively that it reaches human baseline, whereas the same mask does not elevate human thresholds from baseline. In this case, efficiency will approach 100%, but our conclusion should be that human use of the masked information is



(a)



(b)

Figure 4.3: **(A)** Ideal observer RMS contrast thresholds plotted as a function of masking condition. Masking conditions comprise combinations of the left eye (L), right eye (R), nose (N) and mouth (M) as well as a whole-stimulus mask (All). The baseline threshold of 1.7×10^{-6} fell outside the plotting range. **(B)** Log masking, where a value of 0 represents the mean threshold obtained in the low-noise baseline condition. In both figures, error bars representing $\pm$ 1 SEM are plotted but frequently obscured.

highly *suboptimal*. To account for this fact, we transformed efficiency into a measure we term relative efficiency, calculated as:

$$\text{Relative Efficiency} = \frac{\text{masking}_{human}}{\text{masking}_{ideal}} \tag{4.5}$$

where masking is defined in Equation 4.2. On this measure, a value of 1 indicates that human observers suffered the same threshold elevation as the ideal observer, indicating optimal weighting of the masked information. Efficiency and relative efficiency are plotted in Figure 4.4. For the reasons described previously, we focused our remaining analyses on relative efficiency. Specifically, we examined for this measure the difference between horizontally-filtered and vertically-filtered noise masks in each feature condition; any differences between these conditions on this measure indicate differential weighting of the horizontal and vertical bands beyond the extent prescribed by their differential diagnosticity for the task. Descriptive statistics for horizontal selectivity are reported in Table 4.2, and as with raw thresholds we first submitted these data to a 2 (filter) $\times$ 12 (aperture) repeated-measures ANOVA. This analysis revealed significant main effects of filter [$F(1, 14) = 5.53, p = 0.034$], aperture [$F(11, 154) = 43.77, p < 0.0001$], and a significant filter $\times$ aperture interaction [$F(11, 154) = 2.31, p = 0.012$], together demonstrating horizontal selectivity that differs across the face. Next, we submitted all aperture conditions in which both eyes were masked (LR, LRN, LRM, LRNM, All) to a 2 (filter orientation) $\times$ 5 (aperture condition) repeated-measures ANOVA. This analysis revealed significant main effects of filter [$F(1, 14) = 20.23, p = 0.0005$] and aperture [$F(11, 154) = 10.86, p < 0.0001$], but no filter $\times$ aperture interaction [$F(11, 154) = 0.85, p = 0.48$]. This non-significant interaction suggests that more efficient processing of horizontal structure from the eyes

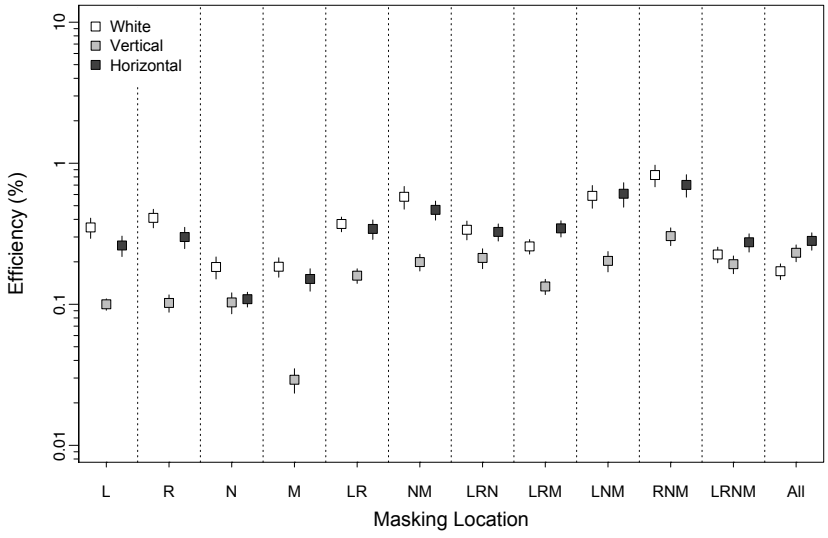
underlied the horizontal selectivity observed when additional features are masked.

	horizontal selectivity			
	μ	t	df	p
L	0.022	1.28	14	0.110
R	0.003	0.28	14	0.393
N	-0.002	-0.17	14	0.567
M	-0.025	-1.33	14	0.898
LR	0.024	2.62	14	0.010
NM	-0.011	-1.31	14	0.894
LRN	0.021	2.03	14	0.031
LRM	0.011	1.41	14	0.090
LNM	-0.010	-0.78	14	0.775
RNM	0.008	0.71	14	0.244
LRNM	0.032	3.09	14	0.004
All	0.028	3.48	14	0.002

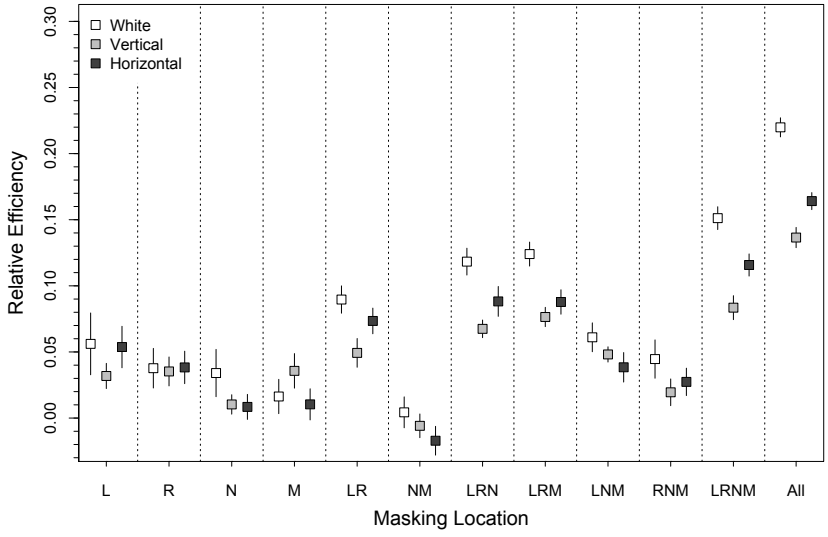
Table 4.2: Paired one-tailed t-tests evaluating the presence of horizontal selectivity in relative efficiency. These tests evaluate the hypothesis that human observers differentially weight the horizontal and vertical band beyond the extent prescribed by their differential diagnosticity for the task.

4.4 Discussion

We measured face identification thresholds for upright faces masked with horizontally-filtered noise, vertically-filtered noise, or white (unfiltered) noise that was passed through Gaussian apertures centred on one or more facial features (left eye, right eye, nose, mouth). Relative to an unmasked baseline, white noise thresholds were maximally elevated when both eyes were masked, demonstrating that observers preferentially relied on these features to perform the identification task. Further, thresholds were elevated more with horizontally-filtered noise than vertically-filtered noise only



(a)



(b)

Figure 4.4: **(A)** Efficiency (%), defined in Equation 4.3. In a masking paradigm such as this, higher efficiency can indicate *less* optimal performance. **(B)** Relative efficiency, defined in equation 4.5. On this measure, a value of 1 indicates equal masking in ideal and human observers, whereas a value of 0 indicates no masking produced by human observers. All error bars represent ± 1 SEM.

when both eyes were masked, demonstrating that observers preferentially relied on the horizontal band when processing information from the eyes. By simulating the performance of an ideal observer under identical conditions, we demonstrated that the eyes carry the most diagnostic information for the identification task, and that the information conveyed by both the eyes and mouth is primarily horizontal. Finally, by comparing human observers to the ideal observer, we demonstrated that human observers make more efficient use of horizontal structure than vertical structure, but only from the eye region.

The present results suggest that selective extraction of information conveyed by horizontal frequency components in the eye region is a signature of upright face identification. Specifically, that the horizontal selectivity observed in studies using whole-face manipulations (Balas *et al.*, 2015; Dakin and Watt, 2009; Goffaux and Dakin, 2010; Goffaux *et al.*, 2011; Jacques *et al.*, 2014; Obermeyer *et al.*, 2012) is likely driven by processes specific to the eyes. Further, given that face identification expertise is associated with both increased horizontal selectivity (Pachai *et al.*, 2013) and more selective attention to the eyes (Heisz and Shore, 2008; Gold *et al.*, 2004; van Belle *et al.*, 2010), the eye selectivity observed in studies using reverse correlation (e.g., Gosselin and Schyns, 2001; Sekuler *et al.*, 2004) and eye tracking (e.g., Arizpe *et al.*, 2012; Blais *et al.*, 2008; Mehoudar *et al.*, 2014; Miellet *et al.*, 2013; Williams and Henderson, 2007) is likely associated with selective extraction of horizontal information from the eyes.

Gold and colleagues (2012) recently examined whether face identification sensitivity can be described using optimal integration of individual feature identification sensitivities (see also Gold *et al.*, 2014; Shen and Palmeri, 2014). In their study, they simulated the performance of an ideal observer, operating under the assumptions of optimal integration, and demonstrated that the eyes contribute maximally to face identification, followed by the nose and, to a slightly lesser extent, the mouth. Our ideal observer analysis reproduced these conclusions precisely (see Figure 4.3). Similarly to the ideal observer, our human observers demonstrated progressively more threshold elevation as additional features were masked (see Figure 4.2). However, our results cannot be compared directly to Gold *et al.* (2012) because our paradigm permitted observers to flexibly utilize unmasked features, whereas their filtering paradigm rendered this strategy impossible. It is likely for this reason we observe minimal effects of masking individual features, and this flexible redistribution of attention also explains why relative efficiency increased as more features were masked, as observers became increasingly unable to avoid the masking noise. Generally, however, our results are consistent with Gold *et al.* (2012), showing that the eyes contribute maximally to overall face identification sensitivity, with the further demonstration that selective attention to information conveyed by horizontal structure underlies this result.

These results may have implications for understanding and ameliorating the deficits experienced by populations with face identification deficits, such as autism (Jiang *et al.*, 2013; Langdell, 1978; Nagai *et al.*, 2013; Rutherford *et al.*, 2007) and prosopagnosia (Barton, 2008; Behrmann *et al.*, 2005; Busigny and Rossion, 2010; Duchaine and Nakayama, 2006). For example, it has been well documented that these groups look less at the eyes during face identification tasks (Barton *et al.*, 2007; Caldara *et al.*,

2005; Nagai *et al.*, 2013; Pelphrey *et al.*, 2002; Schwarzer *et al.*, 2007; Xivry *et al.*, 2008). However, directing attention to the eyes may not ameliorate these deficits, because information from optimal locations may still be processed inefficiently. It remains unclear whether these populations demonstrate decreased sensitivity to the diagnostic horizontal structure in face stimuli, and future research is necessary to address this question. If sensitivity to horizontal structure is correlated with identification performance in these populations, as is the case in typical observers (Pachai *et al.*, 2013), then targeted training regiments could be developed. Such regiments could be particularly well focused if the deficits in processing horizontal structure during face identification are shown to be related to the abnormal patterns of spatial attention commonly observed in such tasks.

Bibliography

Althoff, R. R. and Cohen, N. J. (1999). Eye-Movement-Based Memory Effect: A Reprocessing Effect in Face Perception. *Journal of experimental psychology. Learning, memory, and cognition*, **25**(4), 997–1010.

Arizpe, J., Kravitz, D. J., Yovel, G., and Baker, C. I. (2012). Start position strongly influences fixation patterns during face processing: Difficulties with eye movements as a measure of information use. *PLoS ONE*, **7**(2), 1–17.

Balas, B. J., Schmidt, J., and Saville, A. (2015). A face detection bias for horizontal orientations develops in middle childhood. *Frontiers in Psychology*, **06**, 1–8.

Banks, M. S., Sekuler, A. B., and Anderson, S. J. (1991). Peripheral spatial vision: limits imposed by optics, photoreceptors, and receptor pooling. *Journal of the Optical Society of America. A, Optics and Image Science*, **8**, 1775–1787.

Barton, J. J. S. (2008). Structure and function in acquired prosopagnosia: Lessons from a series of 10 patients with brain damage. *Journal of Neuropsychology*, **2**(1), 197–225.

Barton, J. J. S., Radcliffe, N., Cherkasova, M. V., and Edelman, J. a. (2007). Scan

patterns during the processing of facial identity in prosopagnosia. *Experimental brain research*, **181**(2), 199–211.

Behrmann, M., Avidan, G., Marotta, J. J., and Kimchi, R. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: 1. Behavioral findings. *Journal of cognitive neuroscience*, **17**(7), 1130–49.

Bennett, P. J. and Banks, M. S. (1987). Sensitivity loss in odd-symmetric mechanisms and phase anomalies in peripheral vision. *Nature*, **326**, 873–876.

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., and Caldara, R. (2008). Culture shapes how we look at faces. *PloS one*, **3**(8), e3022.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**(4), 433–6.

Busigny, T. and Rossion, B. (2010). Acquired prosopagnosia abolishes the face inversion effect. *Cortex*, **46**(8), 965–81.

Caldara, R., Schyns, P. G., Mayer, E., Smith, M. L., Gosselin, F., and Rossion, B. (2005). Does prosopagnosia take the eyes out of face representations? Evidence for a defect in representing diagnostic facial information following brain damage. *Journal of cognitive neuroscience*, **17**(10), 1652–66.

Conrey, B. and Gold, J. M. (2009). Pattern recognition in correlated and uncorrelated noise. *Journal of the Optical Society of America. A, Optics, image science, and vision*, **26**(11), B94–109.

Dakin, S. C. and Watt, R. J. (2009). Biological bar codes in human faces. *Journal of Vision*, **9**(4), 1–10.

Duchaine, B. C. and Nakayama, K. (2006). Developmental prosopagnosia: A window to content-specific face processing. *Current opinion in neurobiology*, **16**, 166–173.

Goffaux, V. and Dakin, S. C. (2010). Horizontal information drives the behavioral signatures of face processing. *Frontiers in Perception*, **1**, 1–14.

Goffaux, V., van Zon, J., and Schiltz, C. (2011). The horizontal tuning of face perception relies on the processing of intermediate and high spatial frequencies. *Journal of Vision*, **11**, 1–9.

Gold, J. M., Sekuler, A. B., and Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science*, **28**(2), 167–207.

Gold, J. M., Mundy, P. J., and Tjan, B. S. (2012). The Perception of a Face Is No More Than the Sum of Its Parts. *Psychological science*, **23**(4), 427–434.

Gold, J. M., Barker, J. D., Barr, S., Bittner, J. L., Bratch, A., Bromfield, W. D., Goode, R. A., Jones, M., Lee, D., and Srinath, A. (2014). The perception of a familiar face is no more than the sum of its parts. *Psychonomic bulletin & review*, (2004).

Gosselin, F. and Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision research*, **41**(17), 2261–71.

Heisz, J. J. and Shore, D. I. (2008). More efficient scanning for familiar faces. *Journal of Vision*, **8**(1), 1–10.

Hsiao, J. H.-w. and Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological science*, **19**(10), 998–1006.

Jacques, C., Schiltz, C., and Goffaux, V. (2014). Face perception is tuned to horizontal orientation in the N170 time window. *Journal of Vision*, **14**, 1–18.

Jiang, X., Bollich, A., Cox, P., Hyder, E., James, J., Gowani, S. A., Hadjikhani, N., Blanz, V., Manoach, D. S., Barton, J. J. S., Gaillard, W. D., and Riesenhuber, M. (2013). A quantitative link between face discrimination deficits and neuronal selectivity for faces in autism. *NeuroImage: Clinical*, **2**(1), 320–331.

Langdell, T. (1978). Recognition of faces: an approach to the study of autism. *Journal of Child Psychology and Psychiatry*, **19**(3), 255–68.

Mehoudar, E., Arizpe, J., Baker, C., and Yovel, G. (2014). Faces in the eye of the beholder: Unique and stable eye scanning patterns of individual observers. *Journal of Vision*, **14**, 1–11.

Miellet, S., Vizioli, L., He, L., Zhou, X., and Caldara, R. (2013). Mapping Face Recognition Information Use across Cultures. *Frontiers in psychology*, **4**(February), 34.

Nagai, M., Bennett, P. J., Rutherford, M. D., Gaspar, C. M., Kumada, T., and Sekuler, A. B. (2013). Comparing face processing strategies between typically-developed observers and observers with autism using sub-sampled-pixels presentation in response classification technique. *Vision research*, **79**, 27–35.

Obermeyer, S., Kolling, T., Schaich, A., and Knopf, M. (2012). Differences between Old and Young Adults' Ability to Recognize Human Faces Underlie Processing of Horizontal Information. *Frontiers in aging neuroscience*, **4**(April), 3.

Pachai, M. V., Sekuler, A. B., and Bennett, P. J. (2013). Sensitivity to Information Conveyed by Horizontal Contours is Correlated with Face Identification Accuracy. *Frontiers in Psychology*, **4**(February), 1–9.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, **10**(4), 437–442.

Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., and Piven, J. (2002). Visual Scanning of Faces in Autism. *Journal of Autism and Developmental Disorders*, **32**(4), 249–261.

Peterson, M. F. and Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences*, **109**(48).

Peterson, M. F. and Eckstein, M. P. (2013). Individual Differences in Eye Movements During Face Identification Reflect Observer-Specific Optimal Points of Fixation. *Psychological Science*, **24**(7), 1216–1225.

Rutherford, M. D., Clements, K. A., and Sekuler, A. B. (2007). Differences in discrimination of eye and mouth displacement in autism spectrum disorders. *Vision Research*, **47**(15), 2099–2110.

Schwarzer, G., Huber, S., Grüter, M., Grüter, T., Gross, C., Hipfel, M., and Kennerknecht, I. (2007). Gaze behaviour in hereditary prosopagnosia. *Psychological research*, **71**(5), 583–90.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., and Bennett, P. J. (2004). Inversion leads

to quantitative, not qualitative, changes in face processing. *Current Biology*, **14**(5), 391–6.

Shen, J. and Palmeri, T. J. (2014). The perception of a face can be greater than the sum of its parts. *Psychonomic bulletin & review*, **22**(3), 710–716.

Tanner, W. P. and Birdsall, T. G. (1958). Definitions of d' and eta as psychophysical measures. *The Journal of the Acoustical Society of America*, **30**(10), 922–928.

Tjan, B. S., Braje, W. L., Legge, G. E., and Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research*, **35**(21), 3053–69.

van Belle, G., Ramon, M., Lefèvre, P., and Rossion, B. (2010). Fixation patterns during recognition of personally familiar and unfamiliar faces. *Frontiers in psychology*, **1**, 20.

Watson, A. B. and Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method.

Williams, C. C. and Henderson, J. M. (2007). The face inversion effect is not a consequence of aberrant eye movements. *Memory & cognition*, **35**(8), 1977–1985.

Xivry, J.-J. O., Ramon, M., Lefèvre, P., and Rossion, B. (2008). Reduced fixation on the upper area of personally familiar faces following acquired prosopagnosia. *Journal of Neuropsychology*, **2**(1), 245–268.

Yarbus, A. L. (1967). *Eye Movements During Perception of Complex Objects*. Plenum Press, New York.

Chapter 5

The effect of training with inverted faces on the selective use of horizontal structure

5.1 Introduction

It is commonly believed that most adults are experts at perceiving human faces, and that such expertise is derived, at least in part, from a lifetime of experience with these stimuli (de Heering *et al.*, 2012; Germine *et al.*, 2011; LeGrand *et al.*, 2001; Susilo *et al.*, 2013). Consistent with the claim that perceptual learning contributes to face expertise is the observation that face perception is degraded for less familiar faces, such as those that differ from the observer in terms of age (Malpass and Kravitz, 1969; Rhodes and Anastasi, 2011) or race (Brigham and Barkowitz, 1978; Meissner and Brigham, 2001). Perhaps the most robust demonstration of expertise in face perception is the severe decrement in performance following picture-plane inversion

(Valentine, 1988; Yin, 1969). This face inversion effect (FIE) is particularly interesting because rotation does not alter the physical information available in the stimulus, and therefore the change in performance implies that humans process inverted faces less efficiently than upright faces (Gaspar *et al.*, 2008b; Sekuler *et al.*, 2004).

Previous studies have shown that adults identify faces based on information near the eyes (Gosselin and Schyns, 2001; Sekuler *et al.*, 2004) using a limited band of spatial frequencies (Näsänen, 1999; Gaspar *et al.*, 2008a), and it is plausible to assume that the FIE might be associated with a change in the spatial regions or spatial frequencies that are used to identify inverted faces; however, studies so far have failed to find a significant effect of inversion on the spatial sampling or spatial frequency tuning properties of face identification (Gaspar *et al.*, 2008a; Sekuler *et al.*, 2004; Willenbockel *et al.*, 2011). Recent studies have, however, shown that horizontal facial structure is diagnostic for face identity and expression (Dakin and Watt, 2009; Goffaux and Dakin, 2010; Huynh and Balas, 2014), and that differential sensitivity to this horizontal structure is related to the FIE (Goffaux and Dakin, 2010; Pachai *et al.*, 2013b). These results suggest that face expertise, as indexed by the FIE, is related to changes in the way observers use horizontal structure to identify faces. In this paper, we evaluate this hypothesis by examining whether perceptual learning in an identification task alters sensitivity for horizontal structure in inverted faces.

Many studies have demonstrated that adults can learn to identify upright and inverted faces (e.g., de Heering and Maurer, 2013; Germine *et al.*, 2011; Gold *et al.*, 2004; Hussain *et al.*, 2009a,b; Laguesse *et al.*, 2012). The effects of perceptual learning with faces can be long-lasting (Hussain *et al.*, 2011), and sometimes (e.g., Hussain *et al.*, 2009a, 2011), though not always (e.g. de Heering and Maurer, 2013; Laguesse

et al., 2012) are specific to the trained faces. Training-based improvements in face identification are associated with an increase in calculation efficiency (Gold *et al.*, 1999b, 2004), but it remains unknown how visual processing has been changed to become more efficient. In the present study, we explored the hypothesis that learning enhances observers' ability to extract horizontal structure from faces. To examine this hypothesis, we employed a filtering technique that rendered only a specific band of orientation information diagnostic on a given trial while retaining non-informative power in all other orientation bands. In this way, we were able to measure selective processing of orientation information in perceptually intact faces before and after training. Further, to generalize our results across stimuli, we trained different groups of observers on two different sets of face identities. To normalize our results for differences between these face sets and to characterize the trained improvements with regard to the information available in the stimuli, we compared the results of our human observers to an ideal observer that optimally utilized all available orientation information. Finally, to examine the extent to which trained improvements in identification accuracy transferred to novel exemplars, our observers returned for an additional session in which they were assessed with the untrained face set.

5.2 Methods

5.2.1 Observers

Twenty observers (13 male, $\mu = 23$ years) participated in the experiment. All observers had normal or corrected-to-normal Snellen acuity, and were paid \$10/hour or given partial course credit for their participation. Experimental protocols were

approved by the McMaster University Research Ethics Board, and informed consent was collected prior to the experiment.

5.2.2 Apparatus

The experiment was run on an Apple Macintosh G4 computer using MATLAB and the Psychophysics and Video Toolboxes (Brainard, 1997; Pelli, 1997). Stimuli were displayed on a 21" Apple Studio monitor with a resolution of 1280×1024 (78 pixels/inch) and a frame rate of 85 Hz. An average luminance of 31 cd/m^2 was held constant during the experiment. A chin/head rest was used to stabilize viewing distance at 60 cm, and the experimental apparatus provided the only source of light in the room.

5.2.3 Stimuli

Two sets of ten faces were used in this experiment (Figure 5.1). Both sets were generated using front-facing, digital photographs of five male and five female models. These models had no facial fair, eye glasses, identifiable marks, or visible piercings. Photographs in face set 1 were cropped using a 198×140 pixel oval window and photographs in face set 2 were cropped using a 207×138 pixel oval window. All photographs were centred in a 256×256 pixel matrix, which at the viewing distance of 60 cm subtended $3.7° \times 3.7°$. For more details on the generation of face set 1, see Gold *et al.* (1999a) and for face set 2 see Gaspar *et al.* (2008b).

Across trials, we manipulated the orientation information available to observers by filtering the stimuli in the Fourier domain. Specifically, we selectively retained frequency components from the target face using 18 ideal orientation filters with

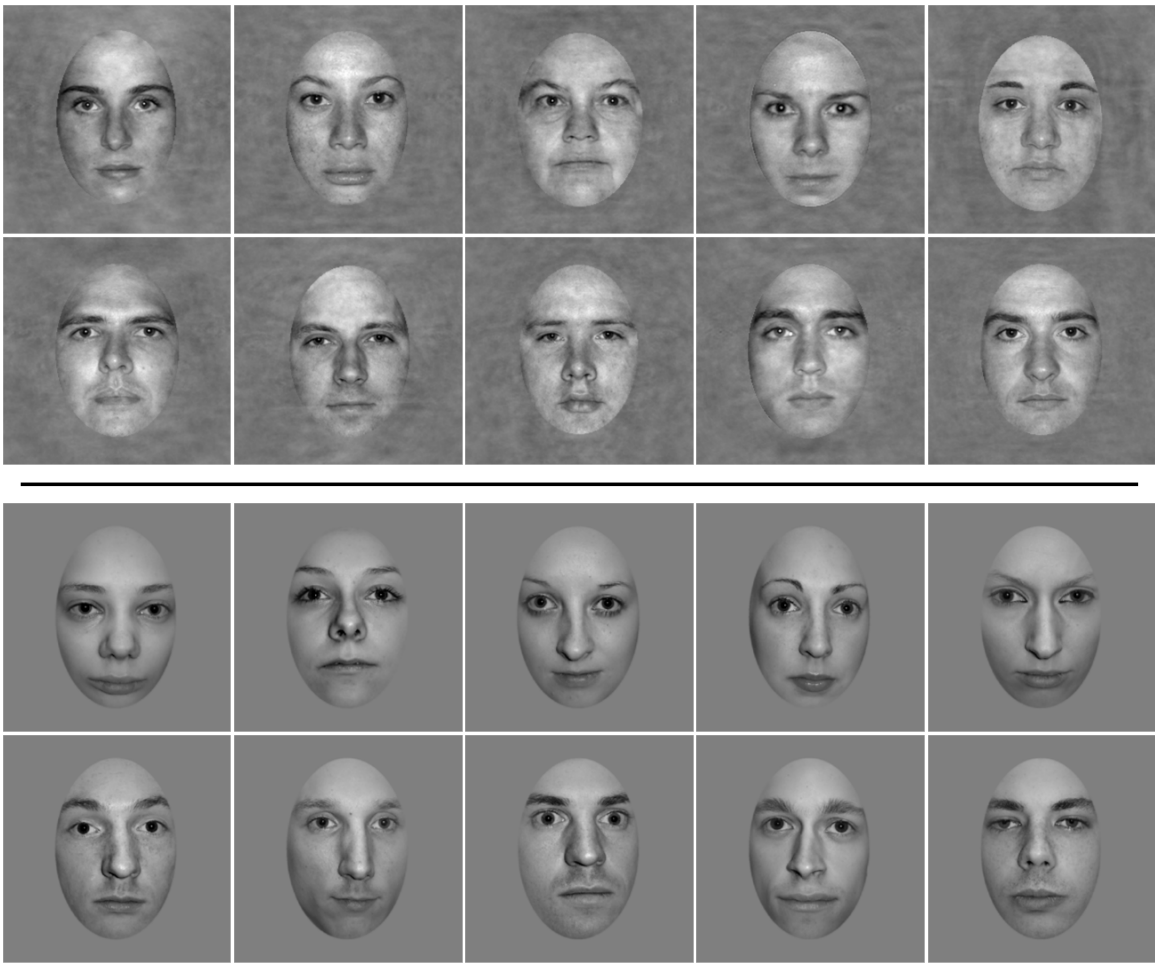


Figure 5.1: Unfiltered examples of the identities included in face set 1 (top) and face set 2 (bottom).

bandwidths ranging from 10° to 180° in 10° steps, centred on 0° (horizontal) or 90° (vertical). Note that 90° is the largest bandwidth at which the horizontal and vertical filters passed independent frequency components, and that 180° filters removed no frequency components and therefore yielded unfiltered faces. Additionally, the components removed from the target face were replaced with components isolated from the average of the 10 possible faces in the relevant face set. Finally, the power of the frequency components from the target face was scaled such that the total power of the combined image was constant across filter conditions. The resulting stimuli contained power in all orientation bands and resembled an unfiltered face, but contained information diagnostic to the identification task in only a limited orientation band. See Figure 5.2 for a demonstration of this process. During the experiment, the stimuli always were inverted by rotating images 180° in the picture plane, and presented at an RMS contrast of 0.5. Stimuli were embedded in white noise with an RMS contrast of 0.1 for later comparison with an ideal observer, which requires external noise.

5.2.4 Design

Observers were randomly assigned to two groups of 10, each trained with a different face set. The training paradigm consisted of a pre-training assessment, three learning sessions, a post-training assessment, and a transfer assessment. Each session was separated by approximately 24 hours except for the transfer session, which was completed approximately 72 hours following the post-training assessment. Pre- and post-training assessments consisted of 10 trials in each condition (2 filter orientations $\times$ 18 bandwidths, 360 total trials), randomly intermixed. Learning sessions each consisted of 300 trials using unfiltered, inverted faces. The transfer assessment consisted

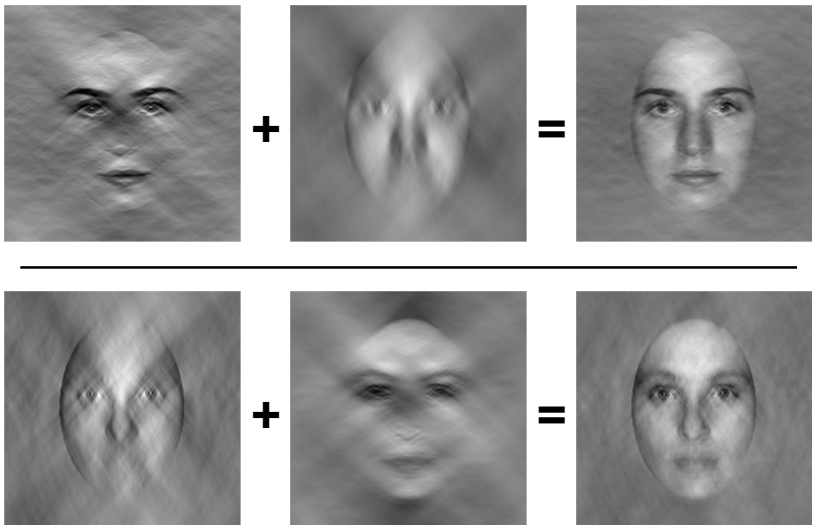


Figure 5.2: Demonstration of the filtering technique using an identity from face set 1. Stimuli on left were passed through a 90° filter centred on horizontal (top) or vertical (bottom). The removed frequency components were replaced with the corresponding components from the average of the 10 faces (centre) to produce the final stimuli presented during the experiment (right).

of two blocks, across which face set was varied. The first block always contained the trained face set, and both blocks were identical in design to the assessment sessions. No practice trials were presented in any of the sessions.

In each session, trials began with a fixation cross presented at the centre of the screen for 500 ms, followed by a 250 ms blank screen. The stimulus then was presented for 500 ms, followed by a 250 ms blank screen and a response window containing front-facing exemplars of the 10 possible faces. Faces in the response window were always unfiltered and inverted. Observers selected their response using a mouse click with no time constraint, and feedback was provided using 600 Hz and 200 Hz tones for correct and incorrect responses, respectively.

5.2.5 Data Analysis

The dependent measure in this experiment was proportion correct in the 10-AFC identification task (P_C). In each session, we fit psychometric functions to P_C as a function of bandwidth for each filter condition using generalized linear models with a probit (i.e., inverse cumulative normal) link function including one additional free parameter, λ, representing the upper asymptote. Given that 90° is the largest bandwidth at which the orientation filters isolated independent frequency components, our measure of horizontal selectivity was the difference in P_C extracted from these psychometric functions at 90° bandwidth for horizontal and vertical filters, respectively.

5.2.6 Ideal Observer Analysis

An ideal observer uses an optimal decision rule to perform a given task, and is limited by the diagnostic information available in the stimulus (Bennett and Banks, 1987; Geisler, 1989, 2011; Gold *et al.*, 2004; Kersten, 1987; Tjan *et al.*, 1995). It has been shown that the ideal decision rule in an identification task will correlate the noisy stimulus on a given trial with templates representing each possible response, then to select the identity that produces the maximum output (Gold *et al.*, 2004; Tjan *et al.*, 1995). Further, when the stimuli are band-pass filtered, the ideal observer bases its decision on only the spatial frequency components carrying diagnostic information, weighing each proportionally to its power. The performance of such an observer can be compared to human observers, quantifying the extent to which they optimally used the available information. This relationship is termed efficiency, and is defined as the squared ratio of human to ideal d' (Tanner and Birdsall, 1958). Therefore, to compute efficiency, human P_C was transformed to d' using the procedures outlined

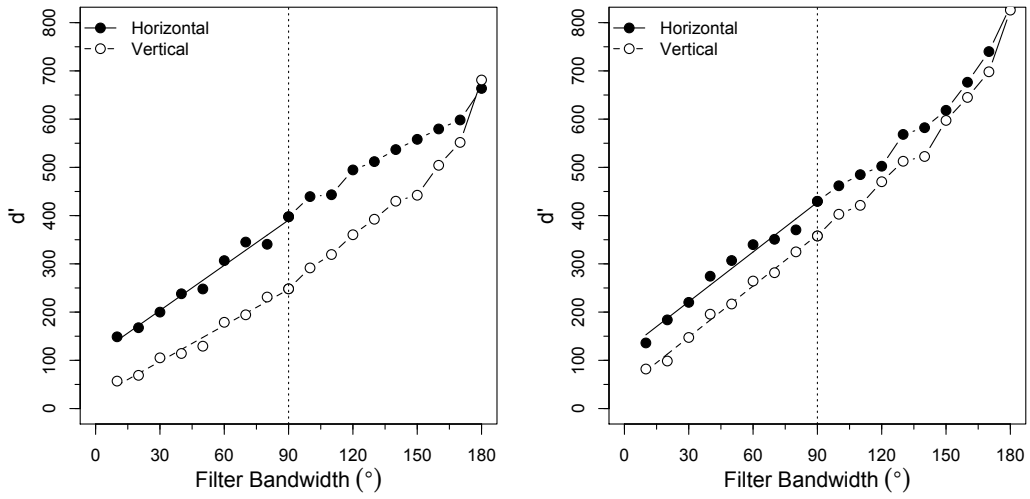
by Macmillan and Creelman (2004). However, at the RMS contrast shown to human observers, the ideal observer would achieve ceiling P_C in all filter conditions, precluding such a transformation. Therefore, for each filter orientation and bandwidth, we simulated ideal performance at seven RMS contrasts ranging from 0.001 to 0.0025, using 5000 trials per contrast level. We then computed best fitting least squares regression lines relating d' to log-transformed RMS contrast, and extrapolated to the contrast presented to human observers. The resulting d' values were used to compute efficiency based on the relationship described above.

5.3 Results

All statistical analyses were conducted using R (R Core Team, 2015). Section 5.3.1 describes the performance of an ideal observer on the 10-AFC identification task, establishing the inherent differences between our two face sets. Section 5.3.2 describes the effect of our training regimen on overall performance and horizontal selectivity. Finally, section 5.3.3 describes the results of our transfer assessment, in which we examined whether trained improvements in performance would generalize to untrained identities. All data for human observers are reported as P_C and efficiency.

5.3.1 Ideal Observer Results

Figure 5.3 plots d' as a function of filter orientation and bandwidth for an ideal observer performing the 10-AFC identification task. The performance of such an observer quantifies the diagnostic information available in the stimulus, and this analysis revealed less total information available in face set 1 than face set 2 [$d'_{180°(set\ 1)} = 672$;



(a) Face Set 1 (b) Face Set 2

Figure 5.3: d' on the 10AFC identification task for an ideal observer using a template matching decision rule that weighs all orientations optimally (see section 5.2.6 for details). Lines represent the best fitting least squares regression fit to the data from 10° to 90° bandwidth, where 90° is indicated by the vertical dotted line.

	Face Set 1			Face Set 2		
Orientation	R^2	Intercept	Slope	R^2	Intercept	Slope
Horizontal	0.98	109.67	3.12	0.98	118.48	3.44
Vertical	0.98	22.87	2.49	0.99	42.16	3.53

Table 5.1: Parameters of the best fitting least-squares regression lines fit to d' as a function of filter bandwidth from 10° to 90° for an ideal observer tested with each face set.

$d'_{180°(set\ 2)} = 830$]. We next examined the differential information available in the horizontal band for each face set by computing best-fitting least squares regression lines relating d' to bandwidth from 10° to 90° for horizontal and vertical filters, respectively. These data were fit well by linear regression (all $R^2 > 0.98$), and the resulting parameters are presented in Table 5.1. This analysis revealed similar slopes between horizontal and vertical filters for both face sets [$\Delta S_{H-V(set\ 1)} = 0.631$; $\Delta S_{H-V(set\ 2)} = -0.097$], which demonstrates that the horizontal advantage is conveyed largely by information in the narrowest bandwidth range. This result is demonstrated further by the large difference between intercepts for horizontal and vertical filters [$\Delta I_{H-V(set\ 1)} = 86.8$; $\Delta I_{H-V(set\ 2)} = 76.3$]. We also observed a smaller difference between intercepts for face set 2, which suggests that the overall informational advantage for this set may be characterized by relatively more information conveyed by the vertical band.

5.3.2 Training Results

Before considering the effect of our training regimen, we examined P_C across the three learning sessions, with each session divided into five 60-trial bins (Figure 5.4). Visual inspection of these data indicates that that much of the learning occurred by the end of the first training session, consistent with previous results on face learning (Gold *et al.*, 1999b, 2004). We verified this observation using a 2 (face set) × 3 (session) × 5 (block) mixed ANOVA with face set as a between-subjects factor. This analysis revealed significant main effects of session [$F(2, 36) = 22.70,\ p < 0.0001$] and block [$F(4, 72) = 15.20,\ p < 0.0001$]. These main effects were qualified by a session × block interaction [$F(8, 144) = 5.12,\ p < 0.0001$]. No other main effects or

interactions approached significance.

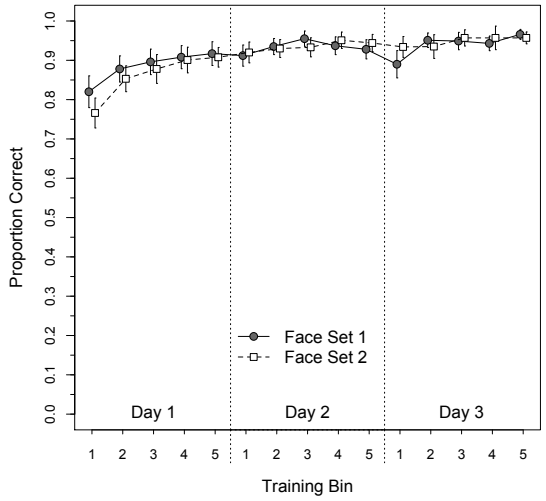
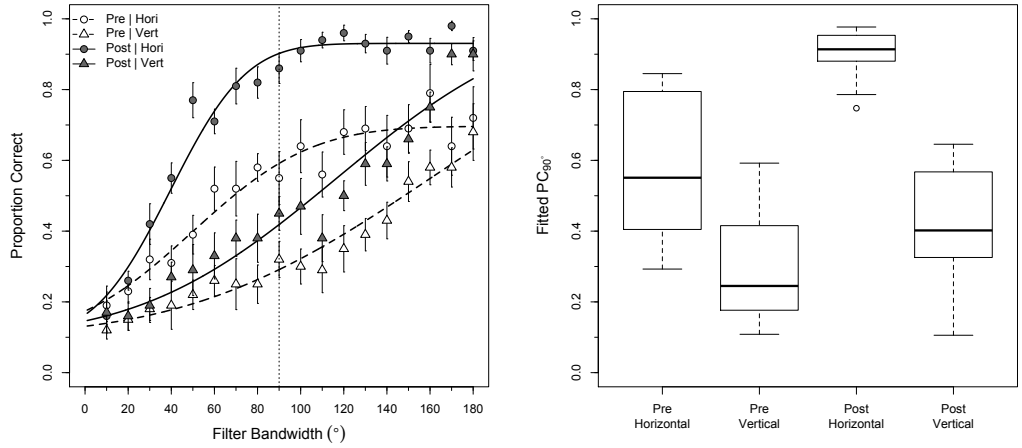


Figure 5.4: Proportion correct during the three learning sessions, divided into 60-trial bins. Error bars represent ± 1 SEM.

Proportion Correct

The effect of our training regimen on P_C for face set 1 is demonstrated in Figure 5.5a. To examine directly the effect of training on horizontal selectivity, we extracted proportion correct at 90° from psychometric functions fit to the entire range of bandwidths and plotted the resulting values in Figure 5.5b. We submitted these data to a 2 (training) × 2 (filter orientation) repeated measures ANOVA, which revealed main effects of training [$F(1,9) = 58.69,\ p < 0.0001$], filter orientation [$F(1,9) = 80.03,\ p < 0.0001$], and a significant interaction [$F(1,9) = 13.34,\ p = 0.0053$]. Together, these results suggest that training improved overall performance (main effect of training) as well as horizontal selectivity (training × filter orientation interaction).

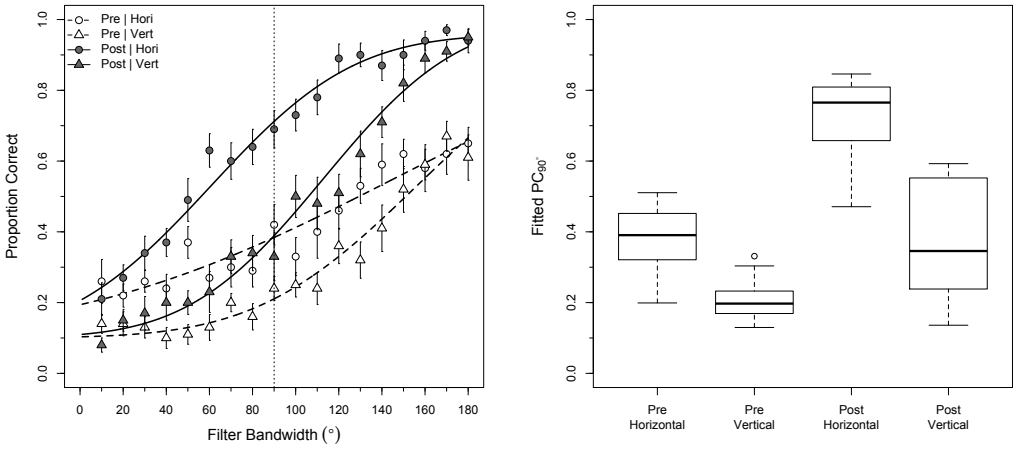
The effect of our training regimen on identification of face set 2 is plotted in Figure



(a) (b)

Figure 5.5: **(a)** Proportion correct on the 10-AFC identification task before and after three days of training with face set 1. Fitted lines are psychometric functions calculated for the mean data. 90° bandwidth is indicated by the vertical dotted line. Error bars represent ± 1 SEM. **(b)** Boxplots of orientation tuning, defined as proportion correct at 90° bandwidth extracted from psychometric functions fit to the entire bandwidth range. The central line represents the median, and the box frame represents the 25th and 75th percentiles, respectively.

5.6. As with face set 1, we extracted proportion correct at 90° from the psychometric functions and submitted the resulting data to a 2 (training) × 2 (filter orientation) repeated-measures ANOVA, which revealed significant main effects of training $[F(1,9) = 48.36,\ p < 0.0001]$, filter orientation $[F(1,9) = 50.96,\ p < 0.0001]$ and a significant interaction $[F(1,9) = 12.35,\ p = 0.0066]$. These results are qualitatively similar to face set 1: training improved overall identification accuracy (main effect of training) as well as horizontal selectivity (training × filter orientation interaction).



(a) (b)

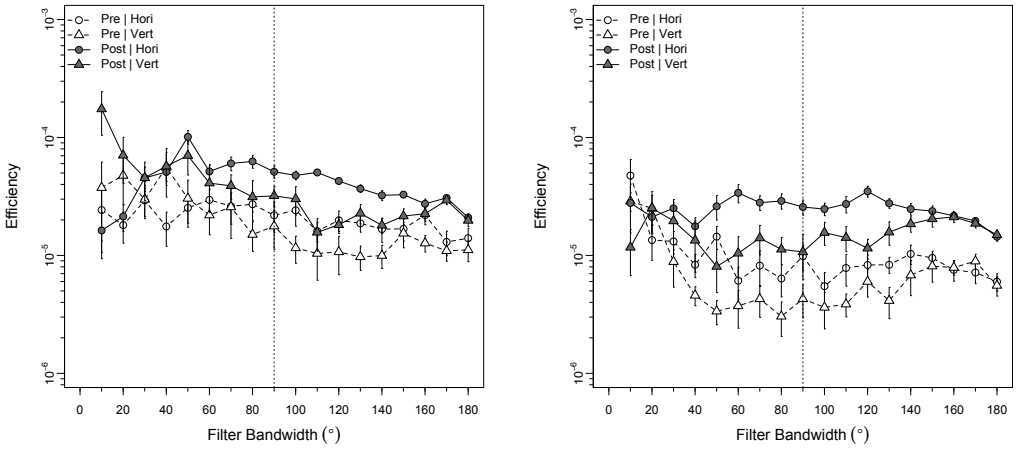
Figure 5.6: **(a)** Proportion correct on the 10-AFC identification task before and after three days of training with face set 2. Fitted lines are psychometric functions calculated using the mean data. 90° bandwidth is indicated by the vertical dotted line. Error bars represent ± 1 SEM. **(b)** Boxplots of orientation tuning, defined as proportion correct at 90° bandwidth extracted from psychometric functions fit to the entire bandwidth range. The central line represents the median, and the box frame represents the 25th and 75th percentiles, respectively.

Efficiency

To rescale the effect of our training regimen with regard to the differential information available in the two face sets, we transformed P_C to efficiency using the relationship described in section 5.2.6. These results are plotted for both face sets in Figure 5.7. We submitted these data, separately for each face set, to 2 (training) $\times$ 2 (filter orientation) $\times$ 18 (bandwidth) repeated-measures ANOVAs.

For face set 1, we observed significant main effects of training [$F(1,9) = 43.15,\ p = 0.0001$] and filter bandwidth [$F(17,153) = 5.01,\ p < 0.0001$] but no significant main effect of filter orientation [$F(1,9) = 0.01,\ p = 0.9213$]. Further, the training $\times$ filter orientation interaction was not significant [$F(1,9) = 0.06,\ p = 0.8111$]. However, the remaining interactions were significant, including training $\times$ bandwidth [$F(17,153) = 1.83,\ p = 0.0289$], filter orientation $\times$ bandwidth [$F(17,153) = 4.61,\ p < 0.0001$], and training $\times$ filter orientation $\times$ bandwidth [$F(17,153) = 2.94,\ p = 0.0002$]. These results differ markedly from P_C in the absence of a training $\times$ filter orientation interaction and the presence of a training $\times$ filter orientation $\times$ bandwidth interaction. This three-way interaction implies that the effect of training on horizontal selectivity differed as a function of bandwidth. Therefore, we divided bandwidth into three equal bins to capture differences in efficiency between narrow bandwidths, near-orthogonal bandwidths, and overlapping bandwidths. These binned data, which are plotted in figure 5.8, were submitted to a 2 (training) $\times$ 2 (filter orientation) repeated-measures ANOVAs separately for each bin. In the first bin (10° to 60°) only the main effect of training reached significance [$F(1,9) = 14.4,\ p = 0.0043$]. In the second bin (70° to 120°) we observed a main effect of training [$F(1,9) = 95.61,\ p < 0.0001$], a main effect of filter orientation [$F(1,9) = 10.17,\ p = 0.011$], and a training $\times$

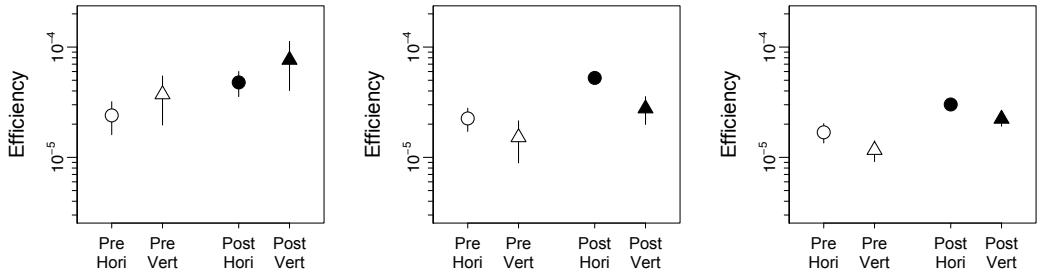
filter orientation interaction [$F(1,9) = 11.44,\ p = 0.0081$], indicating a greater effect of training on efficiency for horizontally filtered faces than for vertically filtered faces. Finally, in the third bin (130° to 180°) we observed main effects of training [$F(1,9) = 118.07,\ p < 0.0001$] and filter orientation [$F(1,9) = 28.88,\ p = 0.0004$], but no training × filter orientation interaction [$F(1,9) = 1.22,\ p = 0.2978$]. Together, these results demonstrate that training enhanced horizontal selectivity for face set 1 (see Figure 5.5), although not at the bandwidth extremes.



(a) Face Set 1 (b) Face Set 2

Figure 5.7: Efficiency relative to an ideal observer that weighs all frequency components optimally before training (open symbols) and after training (closed symbols) for (a) face set 1 and (b) face set 2. See section 5.2.6 for details on efficiency calculation. Error bars represent ± 1 SEM.

For face set 2, we observed significant main effects of training [$F(1,9) = 26.94,\ p = 0.0006$], filter orientation [$F(1,9) = 18.32,\ p = 0.0021$], and bandwidth [$F(17,153) = 4.01,\ p < 0.0001$]. Further, all the two-way interactions were significant, including training × filter orientation [$F(1,9) = 14.67,\ p = 0.004$], training × bandwidth



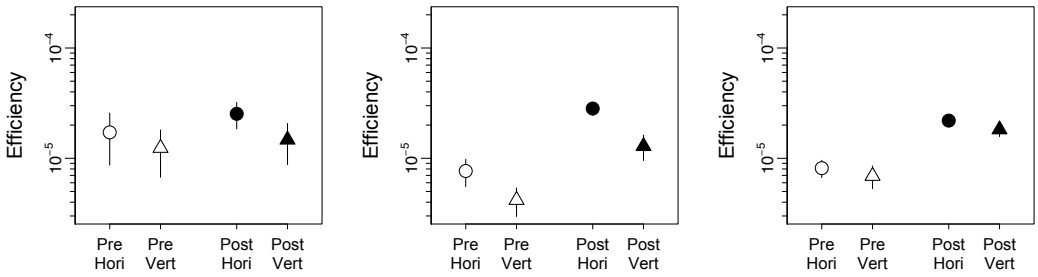
(a) 10° - 60° (b) 70° - 120° (c) 130° - 180°

Figure 5.8: Efficiency relative to an ideal observer that weighs all frequency components optimally before training with face set 1. Data are plotted separately for horizontally and vertically filtered stimuli with narrow, intermediate, or wide bandwidths. Error bars represent ± 1 SEM.

$[F(17, 153) = 4.42,\ p < 0.0001]$, and filter orientation × bandwidth $[F(17, 153) = 2.94,\ p = 0.0002]$. The training × filter orientation × bandwidth interaction was not significant $[F(17, 153) = 0.83,\ p = 0.6613]$, so we did not conduct separate ANOVAs across bandwidth bins, although for comparison purposes we plot the data binned similarly to face set 1 in Figure 5.9. As with face set 1, although training increased overall efficiency for both orientation filtering conditions, the significant training × filter orientation interaction suggests a greater increase in efficiency after training for horizontally filtered faces. Together, these results are qualitatively similar to P_C for face set 2 (see Figure 5.6), demonstrating an increase in overall efficiency and horizontal selectivity after training.

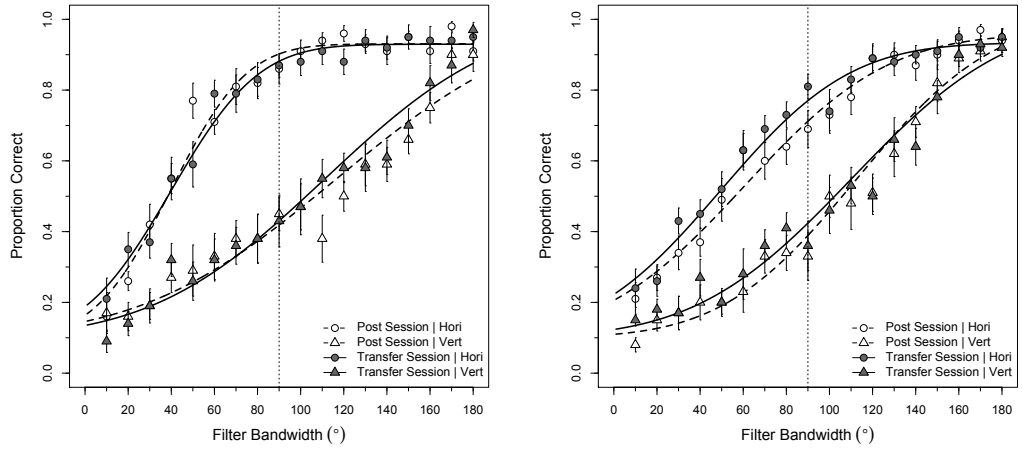
5.3.3 Transfer Results

During the transfer session, observers completed one block with the trained face set and one block with the untrained face set. Figure 5.10 re-plots the post-training data from section 5.3.2 along with proportion correct with trained faces from the transfer



(a) 10° - 60° (b) 70° - 120° (c) 130° - 180°

Figure 5.9: Efficiency relative to an ideal observer that weighs all frequency components optimally before training with face set 2. Data are plotted separately for horizontally and vertically filtered stimuli with narrow, intermediate, or wide bandwidths. Error bars represent ± 1 SEM.



(a) Face Set 1 (b) Face Set 2

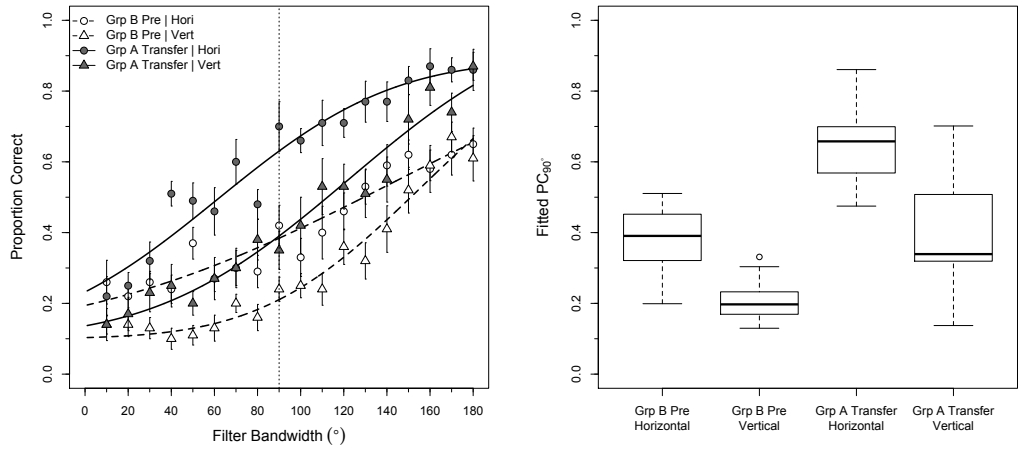
Figure 5.10: Proportion correct on the 10-AFC identification task after training, assessed during the post-training session (open symbols) and the transfer session 72 hours later (closed symbols). The dotted line represents 90° bandwidth, and the error bars represent ± 1 SEM.

session. This comparison demonstrates the extent to which trained improvements degraded in the 72 hours prior to the transfer session. These results clearly show that trained improvements in performance remained intact, so we proceeded to examine transfer to untrained faces.

Proportion Correct

For clarity in the following analyses, we define two groups of observers: group A was trained with face set 1 and group B was trained with face set 2. Figure 5.11a re-plots the pre-training assessment for group B with the transfer assessment for group A. At this point, neither group had any experience with face set 2, so this comparison quantifies the effect of training with face set 1 on face set 2 performance relative to baseline. To examine directly the effect of training on horizontal selectivity, we extracted proportion correct at 90° from psychometric functions fit to the entire range of bandwidths (Figure 5.11b) and submitted the output to a 2 (group) × 2 (filter orientation) mixed ANOVA with group as a between-subjects factor. This analysis revealed significant main effects of group [$F(1, 18) = 25.96,\ p < 0.0001$] and filter orientation [$F(1, 18) = 46.90,\ p < 0.0001$], but no significant interaction [$F(1, 18) = 1.43,\ p = 0.2468$]. These results demonstrate transfer of trained improvements in overall accuracy from face set 1 to face set 2 (main effect of group), although the transfer of horizontal selectivity failed to reach significance (no significant interaction).

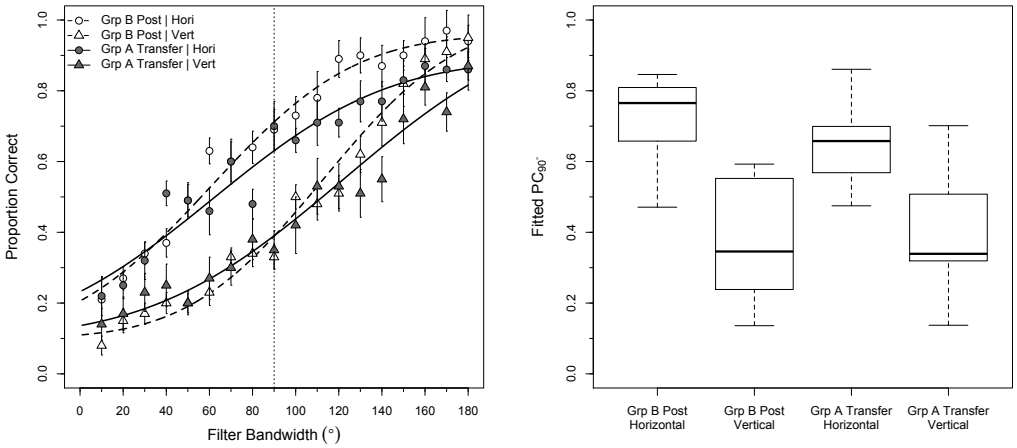
Next, we compared the transfer assessment for group A with the post-training assessment for group B. This comparison, shown in figure 5.12a, quantifies the effect of training with face set 1 on face set 2 performance relative to the effect of training directly with face set 2. We submitted proportion correct at 90° (Figure 5.12b) to a



(a) (b)

Figure 5.11: Proportion correct on the 10-AFC identification task with face set 2. Group A indicates those observers trained with face set 1 and group B indicates those observers trained with face set 2. **(a)** Pre-training assessment for group B (replotted from figure 5.6a) plotted with the transfer assessment group A, following their training with face set 1. Dotted lines represent 90° bandwidth, and all error bars represent ± 1 SEM. **(b)** Boxplots of orientation tuning, defined as proportion correct at 90° bandwidth extracted from psychometric functions fit to the entire bandwidth range. The central line represents the median, and the box frame represents the 25th and 75th percentiles, respectively.

2 (group) $\times$ 2 (filter orientation) mixed ANOVA with group as a between-subjects factor. This analysis revealed a significant main effect of filter orientation [$F(1, 18) = 57.02,\ p < 0.0001$] but no main effect of group [$F(1, 18) = 0.44,\ p = 0.5154$] or interaction [$F(1, 18) = 1.32,\ p = 0.2648$]. These results suggest that training with face set 1 produced comparable effects on identification of face set 2 to training directly with those stimuli.

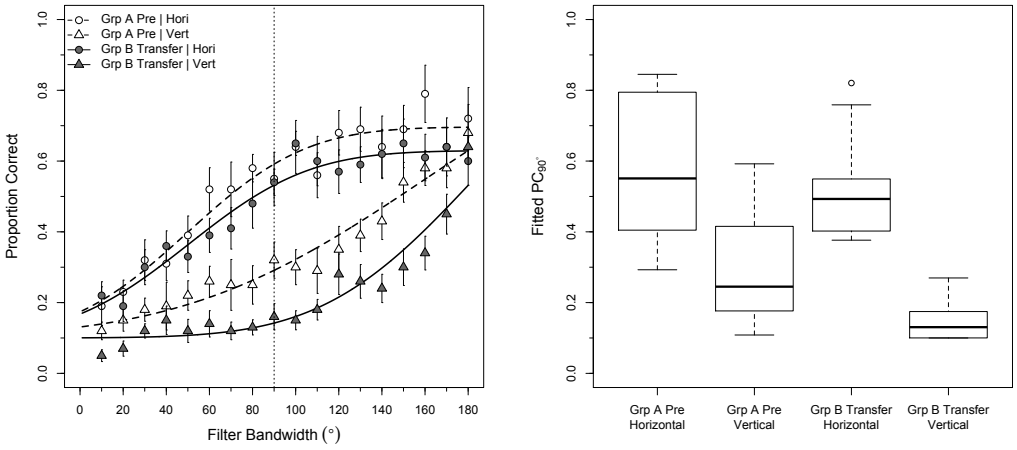


(a) (b)

Figure 5.12: Proportion correct on the 10-AFC identification task with face set 2. Group A indicates those observers trained with face set 1 and group B indicates those observers trained with face set 2. **(a)** Post-training assessment for group B (replotted from figure 5.6a) plotted with the transfer assessment group A, following their training with face set 1. Dotted lines represent 90° bandwidth, and all error bars represent $\pm$ 1 SEM. **(b)** Boxplots of orientation tuning, defined as proportion correct at 90° bandwidth extracted from psychometric functions fit to the entire bandwidth range. The central line represents the median, and the box frame represents the 25th and 75th percentiles, respectively.

Finally, we examined transfer effects in the opposite direction, specifically whether training with face set 2 transferred to face set 1. Figure 5.13a re-plots the pre-training

assessment for group A along with the transfer assessment for group B, comparing two groups of observers with no training on face set 1. A 2 (group) $\times$ 2 (filter orientation) mixed ANOVA on proportion correct at 90° (Figure 5.13b) revealed a significant main effect of filter orientation $[F(1, 18) = 83.96,\ p < 0.0001]$ but no main effect of group $[F(1, 18) = 2.56,\ p = 0.1269]$ or interaction $[F(1, 18) = 1.66,\ p = 0.2146]$. Together, these results demonstrate a failure to transfer overall accuracy (no main effect of group) or horizontal selectivity (no interaction) from face set 2 to face set 1.

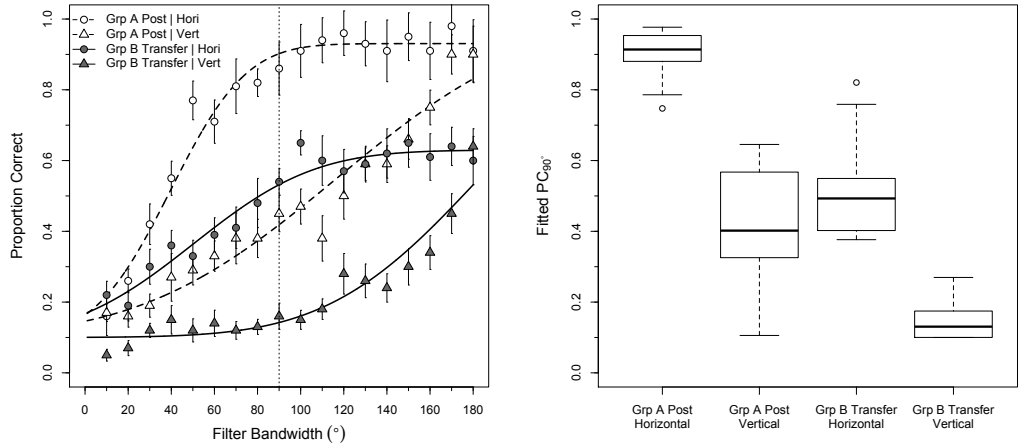


(a) (b)

Figure 5.13: Proportion correct on the 10-AFC identification task with face set 1. Group A indicates those observers trained with face set 1 and group B indicates those observers trained with face set 2. **(a)** Pre-training assessment for group A (replotted from figure 5.5a) plotted with the transfer assessment group B, following their training with face set 2. Dotted lines represent 90° bandwidth, and all error bars represent $\pm$ 1 SEM. **(b)** Boxplots of orientation tuning, defined as proportion correct at 90° bandwidth extracted from psychometric functions fit to the entire bandwidth range. The central line represents the median, and the box frame represents the 25th and 75th percentiles, respectively.

A similar effect was apparent when we compared the transfer assessment for group

B with the post-training assessment for group A. For this comparison, again we submitted proportion correct at 90°, represented in figure 5.14b, to a 2 (group) × 2 (filter orientation) mixed ANOVA, which revealed significant main effects of group $[F(1, 18) = 59.33,\ p < 0.0001]$ and filter orientation $[F(1, 18) = 132.74,\ p < 0.0001]$ but no significant interaction $[F(1, 18) = 2.28,\ p = 0.1482]$



(a) (b)

Figure 5.14: Proportion correct on the 10-AFC identification task with face set 1. Group A indicates those observers trained with face set 1 and group B indicates those observers trained with face set 2. **(a)** Post-training assessment for group A (replotted from figure 5.5a) plotted with the transfer assessment group B, following their training with face set 2. Dotted lines represent 90° bandwidth, and all error bars represent ± 1 SEM. **(b)** Boxplots of orientation tuning, defined as proportion correct at 90° bandwidth extracted from psychometric functions fit to the entire bandwidth range. The central line represents the median, and the box frame represents the 25th and 75th percentiles, respectively.

Together, these results demonstrate asymmetrical transfer between face sets. Specifically, improvements resulting from training with face set 1 transferred to face set 2, but improvements resulting from training with face set 2 did not transfer to set 1.

5.4 Discussion

In the present study, we found that training with inverted faces (300 trials/session for 3 sessions) improved both identification accuracy and selective processing of horizontal structure in the trained stimuli. Further, an ideal observer analysis confirmed that the horizontal band contained more information diagnostic for identity than vertical (see also Pachai *et al.*, 2013b), and that learning increased observers' ability to capitalize on that information difference. In this way, because the observed increase in horizontal selectivity was reflected in both accuracy (Figures 5.5 & 5.6) and efficiency relative to an ideal observer (Figure 5.7), our data demonstrate that training resulted in more efficient sampling of the diagnostic information inherent to face stimuli.

The absolute levels of efficiency in our task were notably lower than those observed for other visual tasks (Banks *et al.*, 1987; Banks and Bennett, 1988; Banks *et al.*, 1991; Bennett *et al.*, 1999; Braje *et al.*, 1995; Gold *et al.*, 2004; Pelli and Farell, 1999; Taylor *et al.*, 2009; Tjan *et al.*, 1995), but were qualitatively similar to the efficiency for 10-AFC face identification measured by Pachai *et al.* (2013b). This low efficiency suggests that sampling of the available stimulus information was highly suboptimal. Such suboptimal sampling is unsurprising, as response classification techniques clearly have established that face identification decisions are based on a spatially limited sample of information from the eyes and eyebrows where stimulus differences are most diagnostic, rather than on the broad range of information spread across the entire face including less diagnostic regions (Gold *et al.*, 2004; Gosselin and Schyns, 2001; Sekuler *et al.*, 2004; Vinette *et al.*, 2004). In fact, Sekuler *et al.* (2004) demonstrated that small differences in spatial sampling are predictive of the face inversion effect. Further, Gold *et al.* (2004) demonstrated that spatial sampling

becomes more efficient with perceptual learning, but that information is selectively sampled from the eye region both before and after training. Finally, recent findings from our lab suggest that observers selectively sample horizontal structure from the eyes during face identification tasks (Pachai *et al.*, 2013a). The current results further elucidate the nature of how stimulus processing improves with training, showing that - even if increased efficiency is constrained to a relatively localized spatial region, as suggested by previous studies - observers improve in their ability to extract structure from the more diagnostic horizontal orientations than from vertical orientations. Taken as a whole, the full set of results across studies suggests that the effects of face inversion and perceptual training reflect modulation of the efficiency with which horizontal structure is sampled from the region around the eyes and eyebrows, with inversion decreasing horizontal selectivity and training increasing it.

Although we observed robust effects of perceptual training, these improvements did not reliably transfer to untrained identities. Specifically, we observed stimulus specificity after training with face set 2, in that trained improvements did not transfer to face set 1, but no stimulus specificity after training with face set 1. In general, stimulus specificity has been the norm in perceptual learning research (Ball and Sekuler, 1987; Fiorentini and Berardi, 1981; Husk *et al.*, 2007; Hussain *et al.*, 2009b, 2012a; Schoups *et al.*, 1995; Yi *et al.*, 2006). However, recent results have suggested that specificity in trained improvements may depend on the training regimen employed. For example, in tasks where learning is typically retinotopic, significant transfer to minimally-trained locations can be induced by intermittent practice at those locations (Xiao *et al.*, 2008; Zhang *et al.*, 2010). Further, Hussain *et al.* (2012b) showed that high inter-stimulus variability while training on an identification task leads to

less stimulus-specificity. Our stimuli had relatively low inter-stimulus variability (one viewpoint, 10 identities), and our improvements did not consistently transfer. Conversely, de Heering and Maurer (2013) demonstrated transfer to novel identities after training on face stimuli with high viewpoint variability. Similarly, Laguesse *et al.* (2012) demonstrated robust transfer after training on a variety of face-related tasks with a large number of identities. Together, these results suggest that training with a larger set of identities may produce a more generalized increase in selective extraction of horizontal structure from face stimuli, although, consistent with Hussain *et al.* (2012b), one might expect more gradual learning with increased stimulus variability.

It is important to note that when designing this experiment, we took care to ensure observers never were cued to the diagnostic band of orientation information. During the training sessions, observers viewed only unfiltered (full-face) stimuli; and during the assessment sessions, we embedded the diagnostic orientation band in a full-face context. For this reason, we can say with confidence that increased face identification accuracy is associated with increased selective sampling of horizontal orientation information. However, it remains unclear whether other experimental paradigms might lead to even greater degrees of learning, and enhanced sensitivity to horizontal orientation structure in the face. For example, paradigms in which observers are trained to utilize the horizontal structure overtly might produce more learning and enhanced horizontal selectivity. Studies aimed at maximizing training-based improvements in face identification have important implications for amelioration of the face identification deficits experienced by a number of populations, including older observers (Grady, 2002; Konar *et al.*, 2013; Maylor and Valentine, 1992; Obermeyer *et al.*, 2012; Rousselet *et al.*, 2009) and individuals with prosopagnosia (Barton *et al.*,

2007; Barton, 2008; Behrmann *et al.*, 2005; Busigny and Rossion, 2010; Duchaine and Nakayama, 2006), schizophrenia (Archer *et al.*, 1992; Christensen *et al.*, 2013; Williams *et al.*, 1999), or autism (Barton *et al.*, 2007; Jiang *et al.*, 2013; Langdell, 1978; Nagai *et al.*, 2013; Rutherford *et al.*, 2007).

5.5 Conclusion

We trained observers to identify intact, inverted faces, measuring identification accuracy and selective processing of horizontal structure before and after training. This study was motivated by demonstrations that horizontal structure is a highly diagnostic cue to face identity (Dakin and Watt, 2009), and that differences in the selective processing of this structure are associated with the magnitude of the FIE (Pachai *et al.*, 2013b). Our results provide further support for the importance of horizontal structure for accurate face identification, as trained improvements in face identification accuracy were associated with more efficient processing of structure within the horizontal orientation band. Although these improvements did not reliably transfer to untrained identities, further studies using modified training regimens may extend the present results to produce, long-lasting generalized improvements in face identification performance.

Bibliography

Archer, J., Hay, D. C., and Young, A. W. (1992). Face processing in psychiatric conditions. *British Journal of Clinical Psychology*, **31**(1), 45–61.

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

Banks, M. S. and Bennett, P. J. (1988). Optical and photoreceptor immaturities limit the spatial and chromatic vision of human neonates. *Journal of the Optical Society of America. A, Optics and image science*, **5**(12), 2059–2079.

Banks, M. S., Geisler, W. S., and Bennett, P. J. (1987). The physical limits of grating visibility. *Vision Research*, **27**(11), 1915–1924.

Banks, M. S., Sekuler, A. B., and Anderson, S. J. (1991). Peripheral spatial vision: limits imposed by optics, photoreceptors, and receptor pooling. *Journal of the Optical Society of America. A, Optics and Image Science*, **8**, 1775–1787.

Barton, J. J. S. (2008). Structure and function in acquired prosopagnosia: Lessons from a series of 10 patients with brain damage. *Journal of Neuropsychology*, **2**(1), 197–225.

Barton, J. J. S., Radcliffe, N., Cherkasova, M. V., and Edelman, J. a. (2007). Scan patterns during the processing of facial identity in prosopagnosia. *Experimental brain research*, **181**(2), 199–211.

Behrmann, M., Avidan, G., Marotta, J. J., and Kimchi, R. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: 1. Behavioral findings. *Journal of cognitive neuroscience*, **17**(7), 1130–49.

Bennett, P. J. and Banks, M. S. (1987). Sensitivity loss in odd-symmetric mechanisms and phase anomalies in peripheral vision. *Nature*, **326**, 873–876.

Bennett, P. J., Sekuler, A. B., and Ozin, L. (1999). Effects of aging on calculation efficiency and equivalent noise. *Journal of the Optical Society of America. A, Optics, image science, and vision*, **16**(3), 654–68.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**(4), 433–6.

Braje, W. L., Tjan, B. S., and Legge, G. E. (1995). Human efficiency for recognizing and detecting low-pass filtered objects. *Vision Research*, **35**(21), 2955–66.

Brigham, J. and Barkowitz, P. (1978). Do "they all look alike?" The Effect of Race, Sex, Experience, and Attitudes on the Ability to Recognize Faces. *Journal of Applied Social Psychology*, **8**(4), 306–318.

Busigny, T. and Rossion, B. (2010). Acquired prosopagnosia abolishes the face inversion effect. *Cortex*, **46**(8), 965–81.

Christensen, B. K., Spencer, J. M. Y., King, J. P., Sekuler, A. B., and Bennett, P. J. (2013). Noise as a mechanism of anomalous face processing among persons with Schizophrenia. *Frontiers in Psychology*, **4**, 1–10.

Dakin, S. C. and Watt, R. J. (2009). Biological bar codes in human faces. *Journal of Vision*, **9**(4), 1–10.

de Heering, A. and Maurer, D. (2013). The effect of spatial frequency on perceptual learning of inverted faces. *Vision Research*, **86**, 107–14.

de Heering, A., Rossion, B., and Maurer, D. (2012). Developmental changes in face recognition during childhood: Evidence from upright and inverted faces. *Cognitive Development*, **27**(1), 17–27.

Duchaine, B. C. and Nakayama, K. (2006). Developmental prosopagnosia: A window to content-specific face processing. *Current opinion in neurobiology*, **16**, 166–173.

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

Gaspar, C. M., Sekuler, A. B., and Bennett, P. J. (2008a). Spatial frequency tuning of upright and inverted face identification. *Vision Research*, **48**(28), 2817–26.

Gaspar, C. M., Bennett, P. J., and Sekuler, A. B. (2008b). The effects of face inversion and contrast-reversal on efficiency and internal noise. *Vision Research*, **48**(8), 1084–95.

Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological review*, **96**(2), 267–314.

Geisler, W. S. (2011). Contributions of ideal observer theory to vision research. *Vision research*, **51**(7), 771–81.

Germine, L. T., Duchaine, B., and Nakayama, K. (2011). Where cognitive development and aging meet: face learning ability peaks after age 30. *Cognition*, **118**, 201–210.

Goffaux, V. and Dakin, S. C. (2010). Horizontal information drives the behavioral signatures of face processing. *Frontiers in Perception*, **1**, 1–14.

Gold, J. M., Bennett, P. J., and Sekuler, A. B. (1999a). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research*, **39**(21), 3537–60.

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

Gold, J. M., Sekuler, A. B., and Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science*, **28**(2), 167–207.

Gosselin, F. and Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision research*, **41**(17), 2261–71.

Grady, C. L. (2002). Age-related differences in face processing: a meta-analysis of three functional neuroimaging experiments. *Canadian Journal of Experimental Psychology*, **56**(3), 208–20.

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

Hussain, Z., Sekuler, A. B., and Bennett, P. J. (2009a). How much practice is needed to produce perceptual learning? *Vision research*, **49**(21), 2624–34.

Hussain, Z., Sekuler, A. B., and Bennett, P. J. (2009b). Perceptual learning modifies inversion effects for faces and textures. *Vision research*, **49**(18), 2273–84.

Hussain, Z., Sekuler, A. B., and Bennett, P. J. (2011). Superior identification of familiar visual patterns a year after learning. *Psychological science*, **22**(6), 724–30.

Hussain, Z., McGraw, P. V., Sekuler, A. B., and Bennett, P. J. (2012a). The Rapid Emergence of Stimulus Specific Perceptual Learning. *Frontiers in psychology*, **3**(July), 226.

Hussain, Z., Bennett, P. J., and Sekuler, A. B. (2012b). Versatile perceptual learning of textures after variable exposures. *Vision Research*, **61**, 89–94.

Huynh, C. M. and Balas, B. (2014). Emotion recognition (sometimes) depends on horizontal orientations. *Attention, perception & psychophysics*.

Jiang, X., Bollich, A., Cox, P., Hyder, E., James, J., Gowani, S. A., Hadjikhani, N., Blanz, V., Manoach, D. S., Barton, J. J. S., Gaillard, W. D., and Riesenhuber, M. (2013). A quantitative link between face discrimination deficits and neuronal selectivity for faces in autism. *NeuroImage: Clinical*, **2**(1), 320–331.

Kersten, D. (1987). Statistical efficiency for the detection of visual noise. *Vision Research*, **27**(6), 1029–1040.

Konar, Y., Bennett, P. J., and Sekuler, A. B. (2013). Effects of aging on face identification and holistic face processing. *Vision research*, **88**, 38–46.

Laguesse, R., Dormal, G., Biervoye, A., Kuefner, D., and Rossion, B. (2012). Extensive visual training in adulthood significantly reduces the face inversion effect. *Journal of Vision*, **12**, 1–13.

Langdell, T. (1978). Recognition of faces: an approach to the study of autism. *Journal of Child Psychology and Psychiatry*, **19**(3), 255–68.

LeGrand, R., Mondloch, C. J., Maurer, D., and Brent, H. P. (2001). Early visual experience and face processing. *Science*, **410**, 890.

Macmillan, N. A. and Creelman, C. D. (2004). *Detection Theory: A User's Guide*. Psychology Press, 2nd edition.

Malpass, R. and Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, **13**(4), 330–334.

Maylor, E. and Valentine, T. (1992). Linear and nonlinear effects of aging on categorizing and naming faces. *Psychology and Aging*, **7**(2), 317–323.

Meissner, C. a. and Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, **7**(1), 3–35.

Nagai, M., Bennett, P. J., Rutherford, M. D., Gaspar, C. M., Kumada, T., and Sekuler, A. B. (2013). Comparing face processing strategies between typically-developed observers and observers with autism using sub-sampled-pixels presentation in response classification technique. *Vision research*, **79**, 27–35.

Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research*, **39**(23), 3824–3833.

Obermeyer, S., Kolling, T., Schaich, A., and Knopf, M. (2012). Differences between Old and Young Adults' Ability to Recognize Human Faces Underlie Processing of Horizontal Information. *Frontiers in aging neuroscience*, **4**(April), 3.

Pachai, M. V., Sekuler, A. B., and Bennett, P. J. (2013a). Masking of individual facial features reveals the use of horizontal structure in the eyes. *Journal of Vision*, **13**(9), 411.

Pachai, M. V., Sekuler, A. B., and Bennett, P. J. (2013b). Sensitivity to Information Conveyed by Horizontal Contours is Correlated with Face Identification Accuracy. *Frontiers in Psychology*, **4**(February), 1–9.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, **10**(4), 437–442.

Pelli, D. G. and Farell, B. (1999). Why use noise? *Journal of the Optical Society of America. A, Optics, image science, and vision*, **16**(3), 647–53.

R Core Team (2015). R: A Language and Environment for Statistical Computing.

Rhodes, M. G. and Anastasi, J. S. (2011). The own-age bias in face recognition: A meta-analytic and theoretical review. *Psychological bulletin*, **138**(1), 146–174.

Rousselet, G. A., Husk, J. S., Pernet, C. R., Gaspar, C. M., Bennett, P. J., and Sekuler, A. B. (2009). Age-related delay in information accrual for faces: evidence from a parametric, single-trial EEG approach. *BMC neuroscience*, **10**, 114.

Rutherford, M. D., Clements, K. A., and Sekuler, A. B. (2007). Differences in discrimination of eye and mouth displacement in autism spectrum disorders. *Vision Research*, **47**(15), 2099–2110.

Schoups, A. A., Vogels, R., and Orban, G. A. (1995). Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. *The Journal of Physiology*, **483**, 797–810.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., and Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, **14**(5), 391–6.

Susilo, T., Germine, L., and Duchaine, B. (2013). Face recognition ability matures late: evidence from individual differences in young adults. *Journal of Experimental Psychology: Human Perception and Performance*, **39**(5), 1212–1217.

Tanner, W. P. and Birdsall, T. G. (1958). Definitions of d' and eta as psychophysical measures. *The Journal of the Acoustical Society of America*, **30**(10), 922–928.

Taylor, C. P., Bennett, P. J., and Sekuler, A. B. (2009). Spatial frequency summation in visual noise. *J Opt Soc Am A Opt Image Sci Vis*, **26**(11), B84–93.

Tjan, B. S., Braje, W. L., Legge, G. E., and Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research*, **35**(21), 3053–69.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, **79**(4), 471–491.

Vinette, C., Gosselin, F., and Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: its in the eyes. *Cognitive Science*, **28**(2), 289–301.

Willenbockel, V., Fiset, D., and Tanaka, J. W. (2011). Relative influences of lightness and facial morphology on perceived race. *Perception*, **40**(5), 621–624.

Williams, L. M., Loughland, C. M., Gordon, E., and Davidson, D. (1999). Visual scanpaths in schizophrenia: is there a deficit in face recognition? *Schizophrenia Research*, **40**(3), 189–99.

Xiao, L.-Q., Zhang, J.-Y., Wang, R., Klein, S. A., Levi, D. M., and Yu, C. (2008). Complete transfer of perceptual learning across retinal locations enabled by double training. *Current biology*, **18**(24), 1922–6.

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

Yin, R. K. (1969). Looking at Upside-Down Faces. *Journal of Experimental Psychology*, **81**(1), 141–145.

Zhang, J.-Y., Zhang, G.-L., Xiao, L.-Q., Klein, S. A., Levi, D. M., and Yu, C. (2010). Rule-based learning explains visual perceptual learning and its specificity and transfer. *The Journal of neuroscience*, **30**(37), 12323–8.

Chapter 6

General Discussion

The goal of this thesis was to characterize a low-level aspect of face stimuli carrying information diagnostic to identity, and to assess whether selective processing of this information is associated with human face identification expertise. In the preceding four chapters, I presented the results of several experiments demonstrating that facial structure conveyed by horizontally-oriented spatial frequency components is diagnostic for face identity, that human observers selectively process this information during face identification tasks, and that this nature of this selective processing is linked to face identification performance. In the following discussion, I will summarize the results of these experiments in more detail, explore their connection to existing research, and suggest several directions for future study.

6.1 Summary of Results

In Chapter 2, I used an ideal observer analysis (Tjan *et al.*, 1995) to demonstrate that horizontal structure is more diagnostic for face identity than the structure conveyed

by any other orientation band. Specifically, I systematically disrupted the information conveyed by each orientation band using filtered noise masks and measured the effect on ideal performance to quantify each band's value for the task at hand. As predicted, ideal identification performance was maximally affected by masking the horizontal band. Second, I showed that human identification performance is characterized by selective processing of horizontal structure. Specifically, using the same manipulations applied to the ideal observer, I measured the relative effect of horizontal masking in human observers along with overall face identification performance. These two measures were significantly correlated, suggesting that efficient use of horizontal structure underlies upright face identification accuracy. Finally, I showed that the face inversion effect is significantly correlated with the loss of horizontal selectivity following inversion. Together, these results were the first to associate face identification expertise with differential use of low-level information.

In Chapter 3, I systematically manipulated filter bandwidth and measured the effect on horizontal selectivity in human and simulated observers. I found that human horizontal selectivity increased with filter bandwidth up to approximately 90°, suggesting that observers integrated across this range of horizontal structure during the identification task. With an ideal observer, I observed approximately constant horizontal selectivity from 10° to 90°, revealing that the most diagnostic information is conveyed by a narrow band ($< 10°$) around the horizontal meridian. However, simulated observers with a fixed bandwidth of approximately 60° - 90°, replicated the pattern of human selectivity more closely. Together, these results suggest that diagnostic horizontal structure is conveyed by a narrow channel, but human observers employ a wider, less optimal orientation sampling strategy to identify faces.

In Chapter 4, I expanded these analyses to the spatial domain. It has long been known that human observers selectively sample information from the eyes and eyebrows during face identification tasks (Gosselin and Schyns, 2001; Heisz and Shore, 2008; Peterson and Eckstein, 2012; Sekuler *et al.*, 2004; Williams and Henderson, 2007; Yarbus, 1967). For this reason, I hypothesized that the horizontal selectivity observed in Chapters 2 and 3 entailed selective extraction of horizontal structure from the eye region. To explore this hypothesis, I used a similar masking technique to Chapter 2, selectively disrupting the information conveyed by horizontal or vertical frequency components. Unlike Chapter 2, I passed these noise masks through one or more Gaussian apertures to selectively disrupt the information conveyed by specific facial features. By measuring the difference between masking horizontal and vertical structure at each location (i.e., left eye, right eye, nose, mouth), I quantified whether observers selectively processed horizontal structure at these locations. Consistent with my hypothesis, observers were more affected by masking horizontal than vertical structure only when the eyes were masked. I also included a series of white noise masking conditions to quantify overall reliance on each feature, and these conditions produced the expected results: observers were more affected by masking the eyes than the nose or mouth. Interestingly, the same manipulations applied to an ideal observer revealed that both the eyes and mouth contain more diagnostic information in the horizontal band. Together, these results demonstrated that human observers base their identity decisions on horizontal structure conveyed near the eyes, despite the availability of other information sources in the face.

Finally, in Chapter 5, I explored whether impaired horizontal selectivity can be improved. Horizontal selectivity is demonstrably reduced when identifying inverted

faces (Goffaux and Dakin, 2010, see also Chapter 2), perceptual learning has been shown to improve inverted face identification (de Heering and Maurer, 2013; Hussain *et al.*, 2009; Laguesse *et al.*, 2012), and face identification accuracy is correlated with horizontal selectivity (Chapter 2). For these reasons, I hypothesized that training with intact, inverted faces would selectively improve the processing of horizontal structure. To assess this hypothesis, I trained observers for three days with inverted faces, assessing horizontal selectivity before and after training using a similar procedure to Chapter 3. I observed, as hypothesized, that training improved sensitivity to horizontal structure and overall inverted face identification accuracy. In fact, comparison of Chapters 3 and 5 reveals that horizontal selectivity and identification accuracy with inverted faces began to resemble that of upright faces after training. Finally, observers returned after 3-4 days to assess whether these trained improvements would transfer to untrained exemplars. Transfer was unreliable (see also Hussain *et al.*, 2009), but the fact that trained improvements in identification accuracy were associated with increased horizontal selectivity is further evidence that efficient use of this information underlies face identification expertise.

6.2 Implications

For decades, the prevailing view has been that upright and inverted faces are processed using qualitatively different mechanisms (see Rossion, 2008 for a review). However, this view has not been universally accepted (Valentine, 1988), and several studies have characterized the face inversion effect as resulting from less efficient, but qualitatively similar processing of inverted faces. For example, Sekuler *et al.* (2004) demonstrated

that observers identify upright and inverted faces in a similar manner, using information from the eyes, and Gaspar *et al.* (2008b) demonstrated directly that face inversion decreases processing efficiency, as opposed to modulating internal noise. Such effects can be understood using hierarchical models of object perception, in which succeeding stages of visual processing involve increasingly complex representations of an object's features (Riesenhuber and Poggio, 1999; Riesenhuber *et al.*, 2004). At the lowest level, units in primary visual cortex are selective for fundamental image features such as orientation (Hubel, 1963). The outputs of these units can be modulated by surrounding stimuli (Allman *et al.*, 1985), or higher-level semantic context (Neri, 2014; Rao and Ballard, 1999). Proceeding along the hierarchy, these outputs are combined to represent contours (Hubel and Wiesel, 1965; Peterhans and von der Heydt, 1989) and eventually three-dimensional shapes (Kourtzi and Kanwisher, 2001) such as faces (Kanwisher and Yovel, 2006). Inversion could affect the efficiency of this pathway by modulating low-level tuning or higher-level feature integration (Riesenhuber and Poggio, 2000).

The effect of inversion on identification performance is one of many interesting phenomena in face perception, and several studies have begun to characterize the role of horizontal structure in these related domains. Goffaux and Dakin (2010) found that horizontal structure is required to observe several signature phenomena in the face perception literature including identity aftereffects, in which adaptation to a single face shifts the perceived identity of a subsequently-presented face away from the adapted identity (e.g., Webster and MacLin, 1999), and composite-face effects, in which information from the bottom-half of an incongruent face affects discrimination of the top half (e.g., Goffaux, 2009; Young *et al.*, 1987). Recently,

Huynh and Balas (2014) found that discrimination of happy and sad facial expressions relies also on horizontal structure. However, horizontal selectivity was reduced when their happy faces included open-mouth smiles, in which the teeth conveyed vertical structure, and was unaffected by rotating the face 90°, such that horizontal structure was conveyed by a vertical channel in veridical coordinates. These findings provide further evidence that horizontal selectivity during face-related tasks is driven by the inherent information content of human faces (see also Chapter 2).

In light of these findings, several results in the face perception literature may be reinterpreted. For example, it is now clear that the spatial frequency range preferred by human observers when identifying faces (8-10 cycles/face, Gaspar *et al.*, 2008a; Willenbockel *et al.*, 2011) is also where the most diagnostic horizontal structure is contained (Goffaux *et al.*, 2011), and that the N170 ERP signature of face processing is largely driven by the information contained in the horizontal band (Jacques *et al.*, 2014). Horizontal selectivity may also capture the role of feature configurations in face identification. Goffaux and colleagues (2007; 2008) have shown that upright face processing is particularly sensitive to vertical relations between facial features. These relations are captured by horizontally-sensitive channels, so this result can also be understood as a decrease in horizontal selectivity following inversion (i.e., Chapters 2 and 3). This account circumvents the definition of a facial feature, which can range from general (i.e., eye) to specific (i.e., iris) with no objective criterion. In fact, several traditional accounts of face perception involving second-order configural properties of facial features (e.g., Diamond and Carey, 1986; Maurer *et al.*, 2002) may be interpreted using orientation-selective models with various-sized receptive fields. For example, Xu *et al.* (2014) outlined recently a computational model that reproduces the

part-whole effect, a classic demonstration of holistic processing in face identification (Tanaka and Farah, 1993). In the model, faces are represented by columns of cells, distributed throughout the visual field, each tuned to numerous spatial scales and orientations. Within each column, those cells with large receptive fields are sensitive to the relative positions of facial features, while those cells with small receptive fields are sensitive to local variations (Xu *et al.*, 2014). With a straightforward modification to include horizontal selectivity, this model would match closely the mechanisms suggested by this thesis, be biologically-plausible, and successfully describe aspects of human behaviour. This is of course just one possible model of face identification, but these characteristics should apply also to future attempts to describe these processes.

6.3 Future Directions

The experiments described in this thesis motivate several directions for future research. For example, in Chapter 2 I demonstrated that sensitivity to horizontal structure predicts face identification performance using a stimulus presentation time of 250 ms. In fact, this is a relatively long time given the impressive speed with which face perception is known to proceed (e.g., Barragan-Jason *et al.*, 2012, 2013; Bieniek *et al.*, 2013; Crouzet and Thorpe, 2011; Jacques and Rossion, 2006), and it remains unclear whether horizontal structure is selectively processed immediately following stimulus presentation. To the extent that object detection precedes discrimination (Or and Wilson, 2010; Pelli *et al.*, 2006; Suchow and Pelli, 2013), rapid horizontal selectivity would imply a role for horizontal structure in face detection. Preliminary results (Pachai *et al.*, 2014) suggest that horizontal structure is extracted from face stimuli after about 100 ms, but further study is required to explore whether this delay

differs across observers or whether these differences have functional consequences for detection and identification performance.

In a typical encounter, human faces convey many cues in addition to identity, and future studies should explore whether horizontal structure underlies these signals. To this end, Huynh and Balas (2014) demonstrated recently that horizontal structure supports discrimination of happy and sad expressions, but it remains unclear whether this is true of other expressions or decisions about age, sex, and intention. If so, it may be optimal for the visual system to automatically process horizontal structure in every face-related scenario. However, if different sources of information support other face-related decisions, an optimal visual system would switch processing strategies for the task at hand. Further study is required to explore whether such strategy switching is necessary, and if so, how successfully it is accomplished.

In Chapter 5, I demonstrated that training with inverted faces improved horizontal selectivity for these stimuli, but this improvement did not reliably transfer to untrained exemplars. This stimulus-specificity may have emerged because observers viewed only one exemplar of each identity. Several recent results suggest that identity coding involves averaging over multiple exposures of the same individual to produce a viewpoint-invariant representation (Burton *et al.*, 2005; de Fockert and Wolfenstein, 2009; Kramer *et al.*, 2015; Leib *et al.*, 2014; Neumann *et al.*, 2013). Indeed, the information conveyed by horizontal structure is also viewpoint-invariant, and these averages may be concisely represented by horizontal structure (Dakin and Watt, 2009). Presenting one exemplar of each face identity may have attenuated this averaging mechanism, encouraging observers to focus instead on image-specific idiosyncrasies. Because horizontal structure is highly diagnostic for identity (Chapter

2), such a focus could improve horizontal selectivity without encouraging a flexible strategy that transfers to untrained exemplars. In fact, presenting more diverse stimuli in a texture identification task does produce more flexible perceptual learning (Hussain *et al.*, 2012), and further studies should explore whether this is true also for learning to process horizontal structure.

These results could have important implications for observers with face identification deficits such as autism (Jiang *et al.*, 2013; Langdell, 1978; Rutherford *et al.*, 2007), prosopagnosia (Barton, 2008; Busigny and Rossion, 2010; Duchaine and Nakayama, 2006) or healthy older observers (Boutet and Faubert, 2006; Habak *et al.*, 2008; Rousselet *et al.*, 2009). Obermeyer *et al.* (2012) recently demonstrated that older observers show less horizontal selectivity when processing faces, but it remains unclear whether this is a general principle that explains the deficits experienced by other populations. Future studies should explore whether these populations demonstrate attenuated horizontal selectivity and whether the extent of this attenuation is associated with the magnitude of their face identification deficits. If so, given also a training regiment that produces flexible improvements in face processing strategies, it may be possible to ameliorate the deficits experienced by these observers.

6.4 Conclusions

Information conveyed by horizontally-oriented spatial frequency components is highly diagnostic for face identity (Chapter 2) and, through a lifetime of experience, human observers are tuned to selectively process this information when making identity judgements. I have shown that this horizontal information is sampled with a relatively

wide bandwidth ($50-90°$, Chapter 3), particularly from the eyes and eyebrows (Chapter 4). This selective horizontal processing is attenuated by face inversion (Chapters 2 and 3), but this attenuation can be ameliorated with perceptual learning (Chapter 5). That horizontal selectivity can be modulated is particularly important because individual differences in this measure are correlated with overall face identification ability (Chapter 2), which suggests opportunities to apply this research to impaired populations. More generally, this research contributes to the growing body of evidence characterizing humans as expert face processors who apply general mechanisms of object perception to detect and categorize face stimuli throughout their lives.

Bibliography

Allman, J., Miezin, F., and McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. *Annual review of neuroscience*, **8**, 407–430.

Barragan-Jason, G., Lachat, F., and Barbeau, E. J. (2012). How fast is famous face recognition? *Frontiers in Psychology*, **3**, 1–11.

Barragan-Jason, G., Besson, G., Ceccaldi, M., and Barbeau, E. J. (2013). Fast and Famous: Looking for the Fastest Speed at Which a Face Can be Recognized. *Frontiers in psychology*, **4**(March), 100.

Barton, J. J. S. (2008). Structure and function in acquired prosopagnosia: Lessons from a series of 10 patients with brain damage. *Journal of Neuropsychology*, **2**(1), 197–225.

Bieniek, M. M., Frei, L. S., and Rousselet, G. a. (2013). Early ERPs to faces: aging, luminance, and individual differences. *Frontiers in psychology*, **4**(May), 268.

Boutet, I. and Faubert, J. (2006). Recognition of faces and complex objects in younger and older adults. *Memory & cognition*, **34**(4), 854–64.

Burton, A. M., Jenkins, R., Hancock, P. J. B., and White, D. (2005). Robust representations for face recognition: The power of averages. *Cognitive Psychology*, **51**(3), 256–284.

Busigny, T. and Rossion, B. (2010). Acquired prosopagnosia abolishes the face inversion effect. *Cortex*, **46**(8), 965–81.

Crouzet, S. M. and Thorpe, S. J. (2011). Low-level cues and ultra-fast face detection. *Frontiers in psychology*, **2**(November), 342.

Dakin, S. C. and Watt, R. J. (2009). Biological bar codes in human faces. *Journal of Vision*, **9**(4), 1–10.

de Fockert, J. and Wolfenstein, C. (2009). Rapid extraction of mean identity from sets of faces. *Quarterly journal of experimental psychology (2006)*, **62**(9), 1716–1722.

de Heering, A. and Maurer, D. (2013). The effect of spatial frequency on perceptual learning of inverted faces. *Vision Research*, **86**, 107–14.

Diamond, R. and Carey, S. (1986). Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology: General*, **115**(2), 107–17.

Duchaine, B. C. and Nakayama, K. (2006). Developmental prosopagnosia: A window to content-specific face processing. *Current opinion in neurobiology*, **16**, 166–173.

Gaspar, C. M., Sekuler, A. B., and Bennett, P. J. (2008a). Spatial frequency tuning of upright and inverted face identification. *Vision Research*, **48**(28), 2817–26.

Gaspar, C. M., Bennett, P. J., and Sekuler, A. B. (2008b). The effects of face inversion

and contrast-reversal on efficiency and internal noise. *Vision Research*, **48**(8), 1084–95.

Goffaux, V. (2008). The horizontal and vertical relations in upright faces are transmitted by different spatial frequency ranges. *Acta psychologica*, **128**(1), 119–26.

Goffaux, V. (2009). Spatial interactions in upright and inverted faces: Re-exploration of spatial scale influence. *Vision Research*, **49**(7), 774–781.

Goffaux, V. and Dakin, S. C. (2010). Horizontal information drives the behavioral signatures of face processing. *Frontiers in Perception*, **1**, 1–14.

Goffaux, V. and Rossion, B. (2007). Face inversion disproportionately impairs the perception of vertical but not horizontal relations between features. *Journal of Experimental Psychology: Human Perception and Performance*, **33**(4), 995–1002.

Goffaux, V., van Zon, J., and Schiltz, C. (2011). The horizontal tuning of face perception relies on the processing of intermediate and high spatial frequencies. *Journal of Vision*, **11**, 1–9.

Gosselin, F. and Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision research*, **41**(17), 2261–71.

Habak, C., Wilkinson, F., and Wilson, H. R. (2008). Aging disrupts the neural transformations that link facial identity across views. *Vision research*, **48**(1), 9–15.

Heisz, J. J. and Shore, D. I. (2008). More efficient scanning for familiar faces. *Journal of Vision*, **8**(1), 1–10.

Hubel, D. H. (1963). The Visual Cortex of the Brain. *Scientific American*, **209**, 54–62.

Hubel, D. H. and Wiesel, T. N. (1965). Receptive Fields and Functional Architecture in Two Nonstriate Visual Areas (18 and 19) of the Cat. *Journal of Neurophysiology*, **28**, 229–289.

Hussain, Z., Sekuler, A. B., and Bennett, P. J. (2009). How much practice is needed to produce perceptual learning? *Vision research*, **49**(21), 2624–34.

Hussain, Z., Bennett, P. J., and Sekuler, A. B. (2012). Versatile perceptual learning of textures after variable exposures. *Vision Research*, **61**, 89–94.

Huynh, C. M. and Balas, B. (2014). Emotion recognition (sometimes) depends on horizontal orientations. *Attention, perception & psychophysics*.

Jacques, C. and Rossion, B. (2006). The time course of visual competition to the presentation of centrally fixated faces. *Journal of Vision*, pages 154–162.

Jacques, C., Schiltz, C., and Goffaux, V. (2014). Face perception is tuned to horizontal orientation in the N170 time window. *Journal of Vision*, **14**, 1–18.

Jiang, X., Bollich, A., Cox, P., Hyder, E., James, J., Gowani, S. A., Hadjikhani, N., Blanz, V., Manoach, D. S., Barton, J. J. S., Gaillard, W. D., and Riesenhuber, M. (2013). A quantitative link between face discrimination deficits and neuronal selectivity for faces in autism. *NeuroImage: Clinical*, **2**(1), 320–331.

Kanwisher, N. and Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **361**(1476), 2109–28.

Kourtzi, Z. and Kanwisher, N. (2001). Representation of Perceived Object Shape by the Human Lateral Occipital Cortex. *Science*, **293**(5534), 1506–1509.

Kramer, R. S. S., Ritchie, K. L., and Burton, A. M. (2015). Viewers extract the mean from images of the same person : A route to face learning. *Journal of Vision*, **15**, 1–9.

Laguesse, R., Dormal, G., Biervoye, A., Kuefner, D., and Rossion, B. (2012). Extensive visual training in adulthood significantly reduces the face inversion effect. *Journal of Vision*, **12**, 1–13.

Langdell, T. (1978). Recognition of faces: an approach to the study of autism. *Journal of Child Psychology and Psychiatry*, **19**(3), 255–68.

Leib, A. Y., Fischer, J., Liu, Y., Qiu, S., Robertson, L., and Whitney, D. (2014). Ensemble crowd perception : A viewpoint-invariant mechanism to represent average crowd identity. *Journal of Vision*, **14**(8), 1–13.

Maurer, D., Grand, R. L., and Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, **6**(6), 255–260.

Neri, P. (2014). Semantic Control of Feature Extraction from Natural Scenes. *Journal of Neuroscience*, **34**(6), 2374–2388.

Neumann, M. F., Schweinberger, S. R., and Burton, a. M. (2013). Viewers extract mean and individual identity from sets of famous faces. *Cognition*, **128**(1), 56–63.

Obermeyer, S., Kolling, T., Schaich, A., and Knopf, M. (2012). Differences between Old and Young Adults' Ability to Recognize Human Faces Underlie Processing of Horizontal Information. *Frontiers in aging neuroscience*, **4**(April), 3.

Or, C. C.-F. and Wilson, H. R. (2010). Face recognition: Are viewpoint and identity processed after face detection? *Vision research*, **50**(16), 1581–1589.

Pachai, M. V., Sekuler, A. B., and Bennett, P. J. (2014). The time course of horizontal tuning during face identification. *Journal of Vision*, **14**(10), 557.

Pelli, D. G., Burns, C. W., Farell, B., and Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, **46**(28), 4646–4674.

Peterhans, E. and von der Heydt, R. (1989). Mechanisms of Contour Perception Contours Bridging Gaps in Monkey Visual Cortex . *Journal of Neuroscience*, **9**(5), 1749–1763.

Peterson, M. F. and Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences*, **109**(48).

Rao, R. P. and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, **2**(1), 79–87.

Riesenhuber, M. and Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature neuroscience*, **2**(11), 1019–1025.

Riesenhuber, M. and Poggio, T. (2000). Models of object recognition. *Nature neuroscience*, **3 Suppl**, 1199–1204.

Riesenhuber, M., Jarudi, I., Gilad, S., and Sinha, P. (2004). Face processing in humans is compatible with a simple shape-based model of vision. *Proceedings of the Royal Society B*, **271 Suppl**, S448–50.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta psychologica*, **128**(2), 274–89.

Rousselet, G. A., Husk, J. S., Pernet, C. R., Gaspar, C. M., Bennett, P. J., and Sekuler, A. B. (2009). Age-related delay in information accrual for faces: evidence from a parametric, single-trial EEG approach. *BMC neuroscience*, **10**, 114.

Rutherford, M. D., Clements, K. A., and Sekuler, A. B. (2007). Differences in discrimination of eye and mouth displacement in autism spectrum disorders. *Vision Research*, **47**(15), 2099–2110.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., and Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, **14**(5), 391–6.

Suchow, J. W. and Pelli, D. G. (2013). Learning to detect and combine the features of an object. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(2), 785–90.

Tanaka, J. W. and Farah, M. J. (1993). Parts and Wholes in Face Recognition. *Journal of Experimental Psychology*, **46**(2), 225–245.

Tjan, B. S., Braje, W. L., Legge, G. E., and Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research*, **35**(21), 3053–69.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, **79**(4), 471–491.

Webster, M. A. and MacLin, O. A. (1999). Figural aftereffects in the perception of faces. *Psychonomic bulletin & review*, **37**(3), 647–653.

Willenbockel, V., Fiset, D., and Tanaka, J. W. (2011). Relative influences of lightness and facial morphology on perceived race. *Perception*, **40**(5), 621–624.

Williams, C. C. and Henderson, J. M. (2007). The face inversion effect is not a consequence of aberrant eye movements. *Memory & cognition*, **35**(8), 1977–1985.

Xu, X., Biederman, I., and Shah, M. P. (2014). A neurocomputational account of the face configural effect. *Journal of Vision*, **14**(8), 1–9.

Yarbus, A. L. (1967). *Eye Movements During Perception of Complex Objects*. Plenum Press, New York.

Young, A. W., Hellawell, D., and Hay, D. C. (1987). Configurational information in face perception. *Perception*, **16**(6), 747–59.