

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Individual Differences in Perception of Mooney Faces

Permalink

<https://escholarship.org/uc/item/4hc997c9>

Author

Canas Bajo, Maria Teresa

Publication Date

2022

Peer reviewed|Thesis/dissertation

Individual Differences in Perception of Mooney Faces

by

Maria Teresa Canas Bajo

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Vision Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor David Whitney, Chair

Professor Michael Silver

Professor William Prinzmetal

Spring 2022

Individual Differences in Perception of Mooney Faces

Copyright 2022
by
Maria Teresa Canas Bajo

Abstract

Individual Differences in Perception of Mooney Faces

by

Maria Teresa Canas Bajo

Doctor of Philosophy in Vision Science

University of California, Berkeley

Professor David Whitney, Chair

Faces play a critical role in our everyday lives. The resulting lifelong experience with faces has made humans remarkable experts in face recognition. One of the main advantages of recognition of faces over other objects is that we process faces as a whole, a face-specific mechanism called holistic perception. Previous work suggests that there is significant variability in face recognition abilities among individuals and across different locations in the visual field. However, the origin of these observer- and location-specific differences remains unclear. In this dissertation, I hypothesized that these differences in face recognition are a result of idiosyncrasies in stimulus-specific holistic processing. To address this, three studies were conducted. First, I tested whether there are individual differences in holistic face recognition and whether these are stimulus-specific. The results indicated that there are substantial stimulus-specific individual differences: the faces that are recognized holistically by one observer are not the same for another observer. Second, I investigated whether holistic face perception is stronger in specific visual field locations and found that sensitivity to faces was relatively tuned toward the fovea. Given the stimulus-specific, observer-specific, and visual-field specific face recognition tuning found in Chapters 2 and 3, I hypothesized that the underlying mechanism may be observer-specific holistic templates, which humans use to detect and recognize faces. The goal of Chapter 4 was to measure and visualize these individual face templates using a classification image approach. The results revealed that there are idiosyncratic differences in face templates that support holistic face recognition and that these are specific to individual observers. Together the results shed light on the existence and nature of individual differences in face recognition, and emphasize the importance of taking into account individual differences in vision research more generally.

Individual Differences; Holistic Perception; Mooney Faces; Face Recognition.

Acknowledgments

When I first considered doing a PhD, many described the process as a solitary one. My experience has been quite the opposite. I would need a whole separate dissertation to properly describe how thankful I am to friends, family and colleagues for the support, the experiences, and the lessons they have provided me across the course of this dissertation, but I hope these lines give them the credit they deserve.

First and foremost, I am immeasurably grateful to my advisor, David Whitney. I am thankful for the countless hours discussing research ideas, dissecting results and figures, and for his support and inspiration to follow the research that I was passionate about. Special thanks to the additional members of my dissertation committee, Michael Silver and Bill Prinzmetal, for their advice and guidance throughout the program. I also want to thank Ken Nakayama and Erv Hafter for their constructive criticism which helped me develop a broader perspective to my thesis.

I am incredibly grateful to have belonged to the Whitney lab for the last five years. Thank you to Zhimin Chen, Mauro Manassi, Yuki Murai, Jefferson Ortega, Zhihang Ren, Zixuan Wang and Allison Yamanashi. I can think of no better lab colleagues.

Being part of the Vision Science program at UC Berkeley has allowed me to cross paths with some of the most incredible women I have ever met. Thank you to Carissa Caloud, for always protecting the students' best interests and to Teresa Puthussery, who gave me the opportunity to learn the art of teaching. Special thanks to Avi Aizenman, Alisa Braun, Kelly Byrne, Vasha Dutell and Michele Winter, for walking this program along with me and for having each other's backs.

Embarking on this academic experience meant moving to the other side of the world from my home country, but I have been very fortunate to create a second home in the Bay Area. I am eternally grateful to my Berkeley family: Angie Godinez, Susan Hao and Sanam Mozaffari. I honestly do not know how I would have survived the writing of this dissertation without these three wise, hilarious, and supporting women. I am also very thankful to my Bay Area Spanish family: Belén Pérez, Cristina Azanza, Hua Chen, Juan Gabriel Fajardo, Juan Vergara, Macarena Abad, Marta Román, Trinidad Puga, and many others, for all the singing, dancing and laughing moments that made me feel right at home.

Even from across the world, I have always relied on the support from my life-long friends back in my hometown: Ana Martín, Alba Orza, David Olmo, Elena González, Irene Arribas, Miguel Malagón, Mireya Serrano, and many others. Thank you for a lifetime of friendship; no matter how far I have been, I always knew I could turn to them for support. Special thanks to my dearest friend Ana Zabala. Words fall short to express the immense gratitude I feel for her friendship.

The most heartfelt thank you to my husband, Arthur. My journey as a Vision Science researcher and as a data scientist literally starts with him teaching me how to code day and night somewhere in Morningside, Edinburgh. There are no words I could write that would appropriately express how grateful and proud I am to have had him by my side throughout my academic training. His unconditional support and belief in me have empowered me to

get through the finish line. He has been a home for me no matter where my research dragged us and no matter how dark it got. For that, I will always be thankful.

Finally, I would like to thank my family. My siblings, Jose, María y Pablo Cañas, whom I love greatly and I'm incredibly proud of. Thank you to Antonio, from whom I have learned the importance of approaching work and relationships with an ethical and empathetic approach, *muchísimas gracias*. Most of all, I want to give the most profound thank you to my mother María Teresa Bajo Molina and my father, Pepe Cañas Delgado. I am incredibly fortunate to have been influenced by their passion to learn, their deep belief in education, their work ethic, and their empathy towards others. It is because I have been a witness to their love for research that I decided to embark on this academic experience. But above it all, I am grateful for their unfailing emotional support. I hope this dissertation makes you proud.

Contents

Contents	iii
1 Introduction	1
2 Stimulus-Specific Individual Differences in Holistic Perception of Mooney Faces	3
2.1 Abstract	3
2.2 Introduction	3
2.3 Experiment 1: Per-face per-observer individual differences.	5
2.3.1 Method	5
2.3.1.1 Participants	5
2.3.1.2 Materials	5
2.3.1.3 Design	6
2.3.1.4 Procedure	6
2.3.1.5 Data Analysis	6
2.3.2 Results	7
2.3.3 Discussion	11
2.4 Experiment 2: an operational definition of Mooney face ‘holistic-ness’.	11
2.4.1 Method	12
2.4.1.1 Participants	12
2.4.1.2 Materials	12
2.4.1.3 Design	12
2.4.1.4 Procedure	13
2.4.2 Results	13
2.4.3 Discussion	15
2.5 General Discussion	17
2.6 Acknowledgments	19
3 Relative Tuning of Holistic Face Processing Towards the Fovea	20
3.1 Abstract	20
3.2 Introduction	21
3.3 Method	24

3.3.1	Participants	24
3.3.2	Material	24
3.3.3	Design	25
3.3.4	Procedure	25
3.3.5	Data analysis	26
3.3.6	Results	26
3.4	Discussion	29
3.5	Acknowledgements	33
4	Individual Differences in Classification Images of Mooney Faces.	34
4.1	Abstract	34
4.2	Introduction	34
4.3	Experiment 1	38
4.3.1	Methods	38
4.3.1.1	Participants	38
4.3.1.2	Material & Design	38
4.3.1.3	Procedure	38
4.3.1.4	Data Analysis	38
4.3.2	Results	40
4.3.3	Discussion	42
4.4	Experiment 2	43
4.4.1	Method	43
4.4.1.1	Participants	43
4.4.1.2	Material	43
4.4.1.3	Design, procedure and data analysis	43
4.4.2	Results	44
4.4.3	Discussion	47
4.5	General discussion	48
4.6	Acknowledgments	50
5	Conclusion	51
6	Bibliography	53

Chapter 1

Introduction

Face recognition is critical to most aspects of our everyday lives. They are center of most social interactions, and they provide meaningful emotional and non-verbal information. The crucial social function of faces and the prolonged exposure to them has led to the development of advantageous face-specific mechanisms. Humans prefer faces over other objects since early infancy (Farzin, Rivera & Whitney, 2009; Frank, Vul & Johnson, 2009; Umiltà, Simion & Valenza, 1996), and are able to quickly find and identify faces in the visual world (Bindemann, Burton, Hooge, Jenkins & de Haan, 2005; Cerf, Frady & Koch, 2009; Cerf, Harel, Einhauser & Koch, 2008; Costela & Woods, 2019; Coutrot & Guyader, 2014; Foulsham, Cheng, Tracy, Henrich & Kingstone, 2010; Jack & Schyns, 2015; Marat, Rahman, Pellerin, Guyader & Houzet, 2013; Ro, Russell & Lavie, 2001; Theeuwes & Van der Stigchel, 2006). In fact, humans can recognize faces under a wide range of conditions, including under a challenging environment, (Burton, Bruce & Hancock, 1999; Hole, George, Eaves & Rasek, 2002) where artificial face recognition models face difficulties (Phillips & O'Toole, 2014).

One of the most critical advantages in how we perceive faces compared to objects is that while we perceive objects by processing each face feature part by part (an eye, a nose, a mouth), we perceive faces holistically. That is, we perceive the face as a whole rather than a separate set of face features (Sergent, 1984). Evidence for holistic processing of faces and its advantages is extensive (Farah, Wilson, Drain & Tanaka, 1998; Maurer, Le Grand & Mondloch, 2002), the strongest of them being the inversion effect (Yin, 1969). This phenomenon describes the fact that upright faces are more efficiently processed than upside-down faces. The advantage of upright faces comes from our ability to process them holistically. Holistic processing, in contrast, is broken in upside-down faces, which makes them harder to identify (Busey & Vanderkolk, 2005; McKone, 2004). Humans can still recognize upside-down separate face features without problems since part-based processing is not affected by the inversion (McKone, 2004). Further evidence for holistic processing of upright faces comes from Mooney faces. Mooney faces are two-tone images that are readily perceived as faces even though they lack low-level face features (Mooney, 1957). Mooney faces require holistic processing to be recognized as faces, and therefore upside-down are difficult or even impossible to recognize (Kanwisher, Tong & Nakayama, 1998; Latinus &

Taylor, 2005; McKone, 2004). Perception of Mooney faces is evidence that we process faces as a whole, and for that reason they are often used to study holistic perception.

Previous work suggests that face recognition is supported by matching observed faces to stored templates (Cavanagh, 1991; Rossion, 2008; Rossion & Boremanse, 2008; Valentine, 1991). Faces most similar to those internal face representations are more efficiently processed than that most dissimilar faces. Indeed, previous research in face perception suggests that recognition of faces is tuned to faces that we encounter most often (Dobs, Isik, Pantazis & Kanwisher, 2019; Landi & Freiwald, 2017; Ramon, Vizioli, Liu-Shuang, & Rossion, 2015; Visconti di Oleggio Castello, Halchenko, Guntupalli, Gors., & Gobbini, 2017). For instance, we only process upright faces holistically, which are the most common orientation faces appear at in our visual environment. Contact with upside-down faces is very limited so development of face recognition mechanisms has not been tuned towards them. This work suggests that internal face templates are tuned to preferred faces at the group-level. It also raises the question of whether internal face templates and by consequence face recognition are tuned also at the observer and stimuli level.

In the series of studies presented here, we use Mooney faces to achieve three goals. First, Chapter 2 aims to investigate whether there are stimuli-specific individual differences in holistic processing of Mooney faces. Second, Chapter 3 explores whether face recognition is also visual-field-specific. That is, whether face recognition is tuned to specific visual field locations. In Chapter 4, our goal is to investigate whether idiosyncrasies in face recognition reflect underlying individual differences in internal face representations – templates.

Chapter 2

Stimulus-Specific Individual Differences in Holistic Perception of Mooney Faces

2.1 Abstract

Humans perceive faces holistically rather than as a set of separate features. Previous work demonstrates that some individuals are better at this holistic type of processing than others. Here, we show that there are unique individual differences in holistic processing of specific Mooney faces. We operationalized the increased difficulty of recognizing a face when inverted compared to upright as a measure of the degree to which individual Mooney faces were processed holistically by individual observers. Our results show that Mooney faces vary considerably in the extent to which they tap into holistic processing; some Mooney faces require holistic processing more than others. Importantly, there is little between-subject agreement about which faces are processed holistically; specific faces that are processed holistically by one observer are not by other observers. Essentially, what counts as holistic for one person is unique to that particular observer. Interestingly, we found that the per-face, per-observer differences in face discrimination only occurred for harder Mooney faces that required relatively more holistic processing. These findings suggest that holistic processing of hard Mooney faces depends on a particular observer's experience whereas processing of easier, cartoon-like Mooney faces can proceed universally for everyone. Future work using Mooney faces in perception research should take these stimulus-specific individual differences into account to best isolate holistic processing.

2.2 Introduction

Humans perceive faces holistically rather than as a set of separate features (Sergent, 1984), and it is our ability to perceive faces as a whole that makes humans experts in face processing

(Farah et al., 1998; Maurer et al., 2002). One of the strongest pieces of evidence for holistic perception of faces is the inversion effect (Yin, 1969); holistic processing breaks down with inverted faces (Farah, Tanaka & Drain, 1995; Rossion, 2008; Sargent, 1984). This results in reduced performance (lower accuracies and longer reaction times) when identifying inverted faces compared to upright faces (Busey & Vanderkolk, 2005; McKone, 2004; Yin, 1969), but there is no such drop in accuracy for inverted single features (McKone, 2004). Metrics like the composite-face test (McKone, 2008) and the whole-part test (Donnelly & Davidoff, 1999) provide further empirical evidence of holistic processing of faces.

Studies of holistic processing often use the inversion effect as an operational measure of ‘holistic-ness’ (McKone, 2004; Taubert, Apthorp, Aagten-Murphy & Alais, 2011) and frequently use Mooney faces as stimuli (Andrews & Schluppeck, 2004; George, Jemel, Fiori, Chaby & Renault, 2005; Latinus & Taylor, 2005; Leo & Simion, 2009; Verhallen et al., 2014). Mooney faces (Fig. 2.1) are two-tone black and white blobs that are readily perceived as faces despite lacking low-level segmentable face-specific features (Mooney, 1957). Mooney face recognition is all-or-none, either you perceive it as a face in a glance or it remains just an abstract impression of blobs (Brodski, Paasch, Helbling & Wibrall, 2015; Schwiedrzik, Melloni & Schurger, 2018), much like Dallenbach’s cow (Dallenbach, 1951) and Gregory’s Dalmatian dog (Gregory, 1970). Inversion of a Mooney face makes its identification almost impossible (Kanwisher et al., 1998; Latinus & Taylor, 2005; McKone, 2004). Because the inversion effect is stronger for Mooney faces than for gray-scale faces (McKone 2004), they are often treated as a more efficient means of isolating holistic face processing.

There are a large variety of Mooney faces that have been created and published, including the original artist-created images (Mooney, 1957), thresholded images of gray-scale faces (Brodski et al., 2015; Goold & Meng, 2016; Schwiedrzik et al., 2018), and artificially created images based on machine learning approaches (Ke, Yu & Whitney, 2017). While Mooney faces are typically treated as an excellent stimulus to isolate holistic processing, there are also significant variations between different Mooney faces (Fig. 1). This raises the question of whether different Mooney faces may be more or less effective at isolating the holistic mechanisms of face recognition.

Here, we measured holistic processing of individual Mooney faces using the magnitude of the inversion effect, a method widely accepted as a marker of holistic processing (McKone & Yovel, 2009; Taubert et al., 2011; Yin, 1969). Unlike previous studies, our approach measures the strength of the inversion effect for individual Mooney faces to identify the magnitude of holistic processing for specific stimuli and for specific observers. If there are consistent face-specific differences, it would suggest that individual face images should be evaluated before assuming that they tap into “holistic” processing. Additionally, there may be individual subject differences that interact with the individual faces: the particular faces that are holistic for one subject may or may not be the same as those for another subject. Assessing holistic processing within any given observer may require tailoring the stimuli to that subject’s unique fingerprint of face processing.

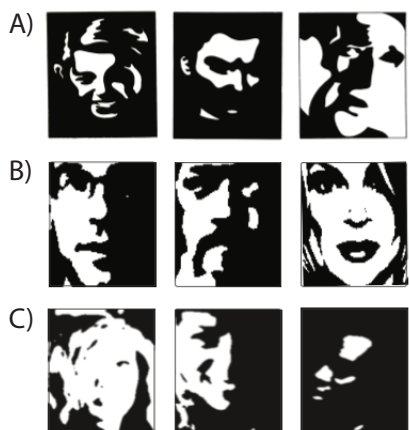


Figure 2.1: Example of three Mooney face datasets. (A) Mooney faces adapted from Mooney (1957). (B) Mooney faces created by a generative adversarial network model adapted from Ke et al. (2017). (C) Mooney faces created from thresholds of grayscale faces adapted from Schwiedrzik et al., 2018.

2.3 Experiment 1: Per-face per-observer individual differences.

2.3.1 Method

2.3.1.1 Participants

Thirty participants took part in this experiment. All subjects were undergraduate students at the University of California, Berkeley and provided written consent form before participation. All experimental procedures were approved by the UC Berkeley Institutional Review Board.

2.3.1.2 Materials

The stimuli consisted of 192 images (96 ambiguous Mooney faces and 96 shuffled Mooney stimuli). Literally thousands of Mooney faces have been created and published in the past (Brodski et al., 2015; Goold & Meng, 2016; Ke et al., 2017; Schwiedrzik, et al., 2018). In the present study, we used a subset of Mooney faces created by Schwiedrzik et al. (2018) because these Mooney faces have been used to study individual differences, and because these stimuli include a diverse array of faces (varying in perspective, lighting direction, face identity, gender, etc.). It is worth noting that this data set is unique in that the scrambled stimuli were created manually from each respective Mooney face so it would preserve the main structure of both upright and inverted versions. The scrambled stimuli were modified in an iterative process in which each shuffled stimulus was tested separately until 85% of the

subjects agreed that there was ‘no face’ in it (see Schwiedrzik et al., 2008 for more controls and details). Each Mooney face had only one scrambled stimulus, so that the pair always appeared together. A CRT monitor at 60Hz was used to display all the stimuli. The monitor was placed at 60cm from the participant and at this distance Mooney stimuli subtended a visual angle of 10 degrees. A chin rest was used to stabilize subjects head and distance to the monitor. The presentation of the stimuli was controlled using MATLAB R2016b with Psychophysics Toolbox 3 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007).

2.3.1.3 Design

Each subject completed a total of 1920 trials (2 (Upright vs. Inverted) x 96 Mooney stimuli x 10 repetitions of each Mooney Face). The faces were repeated ten times in each condition (upright and inverted) to generate enough trials to facilitate a split-half within subjects’ correlation procedure and increase the power of its results.

2.3.1.4 Procedure

The task was to report which of two images was a face (left or right). Participants were instructed to ignore the orientation of the face (i.e., an upside-down face could be the correct answer). Each trial started with a fixation cross in the center of the screen displayed for one second. Immediately after, two Mooney images were displayed, a Mooney face and a scrambled Mooney stimulus, one to the left of the fixation cross and the other one to the right of the fixation cross (Fig. 2). The location of each stimulus was randomly chosen to be left or right in each trial. The orientation of the Mooney face was upright on half of the trials and inverted on the other half. The two conditions were randomly interleaved between trials. The two images were shown until a response was given (self-paced). Subjects were instructed to give their answers as fast as possible. After the stimuli onset, participants were allowed to free-view the two images.

2.3.1.5 Data Analysis

The magnitude of the inversion effect for each Mooney face was calculated per subject by subtracting the average accuracy of that Mooney face in the upright condition and the average accuracy of that Mooney face in the inverted condition.

For all analyses, we used a bootstrap re-sampling procedure (Efron & Tibshirani, 1994). In each iteration of this procedure, the trial information used to calculate the average for upright and inverted conditions per subject, per Mooney face, was re-sampled with replacement for only five of the ten trials available for each subject, face and condition. The reason to only choose half of the trials was to allow for a split-half procedure to quantify within-subject consistency. To generate the null distributions for the within- and between-subject correlations, we shuffled the label of the Mooney face number and then bootstrapped the correlation without replacement.

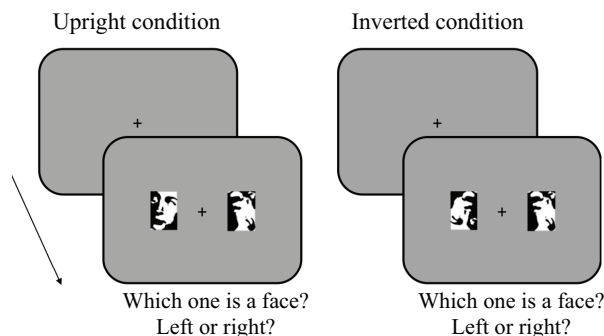


Figure 2.2: Task in Experiment 1 for the upright condition (left) and inverted condition (right). In both conditions, subjects had to first fixate a cross at the center of the screen. Then, two images appeared at either side of the screen, left or right. The observers' task was to discriminate which of the two images was a face, by pressing the left or right arrow.

2.3.2 Results

Our initial goal was to quantify the per-subject individual differences in holistic processing of Mooney faces. We operationalized the extent to which a subject processed Mooney faces holistically as the average magnitude of the inversion effect (average accuracy in the upright condition minus the average accuracy in the inverted condition) across all Mooney faces (Fig. 3). We found that the magnitude of the inversion effect varied significantly across subjects (Fig. 3B). The individual differences found here replicate prior work (Russel, Duchaine & Nakayama, 2009; Wang, Li, Fang, Tian & Liu, 2012).

Our primary goal was to investigate the per-face differences in holistic perception for individual subjects. That is, are the individual differences stimulus specific? To address this, we calculated the magnitude of the inversion effect for each Mooney face across subjects. We found that individual Mooney faces varied significantly in the magnitude of their inversion effect (Fig. 4B) and these differences reflected changes in performance in both upright and inverted conditions (Fig. 4A). That is, not all Mooney faces tapped into holistic processing to the same extent.

As a simple confirmation of the previously documented inversion effect (Yin, 1969), we also calculated the average accuracy collapsed across Mooney faces and subjects for the upright and inverted conditions, separately. Replicating the extensively studied inversion effect, we found higher accuracy for upright (Mean = 0.90; SD = 0.016) than for inverted faces (Mean = 0.742; SD = 0.028) across all subjects and Mooney faces, $t(95) = 9.618$, $p < .001$ (Fig. 4C).

In addition, we tested whether subjects were consistent in recognizing the Mooney faces across their 10 trials (Fig. 4D). We found that if recognized in the first instance, the same

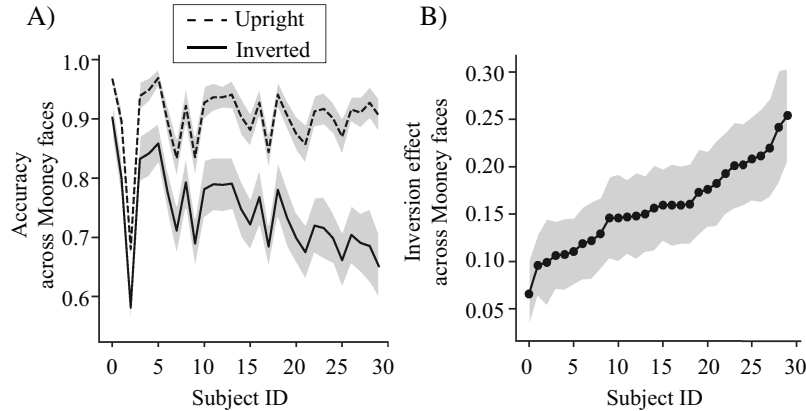


Figure 2.3: Results of Experiment 1. (A) Upright Mooney face (dashed) and inverted Mooney face (solid) discrimination accuracy (ordinate), averaged across all Mooney faces, for each subject (abscissa). The shaded region represents the 95% CIs around the mean performance for each subject. (B) Magnitude of the inversion effect (the difference between average accuracy in the upright condition and average accuracy in the inverted condition) for each subject (abscissa). Participants (subject ID) are ordered from those with weaker inversion effects to those with stronger inversion effects. The shaded region represents the 95% CIs around the mean inversion effect for each subject.

upright Mooney face image was also recognized in subsequent presentations as well (Fig. 4D, slope parameter of linear regression $b = -0.07$, $p = 0.05$). Conversely, if the Mooney face was not recognized in the first trial, it was also not recognized in subsequent trials. Subjects did get slightly better at recognizing inverted Mooney faces, but the effect was modest (Fig. 4C; slope parameter of linear regression, $b = 0.11$, $p < 0.05$).

So far, we replicated the extensive literature showing that subjects differed in the extent they process faces holistically and we extended these results by showing that not all Mooney faces tap into holistic processing to the same extent. Next, we wanted to investigate whether subjects agreed on which Mooney faces are processed holistically. To that end, we investigated the between-subject agreement in the magnitude of the inversion effect for each Mooney face by calculating the mean average pairwise correlations of subjects' per-face inversion effect. We found relatively low between-subject correlation in the inversion effect (Fig 5C; $r = 0.106$, $CI = [0.085, 0.127]$), although it remained significantly different from zero ($p < 0.001$). Interestingly, the low between-subject agreement reveals that subjects do not process Mooney faces holistically to the same extent (Fig. 5C). In fact, we found that there was a low between-subject correlation in both upright and inverted accuracies (Fig. 5A-B).

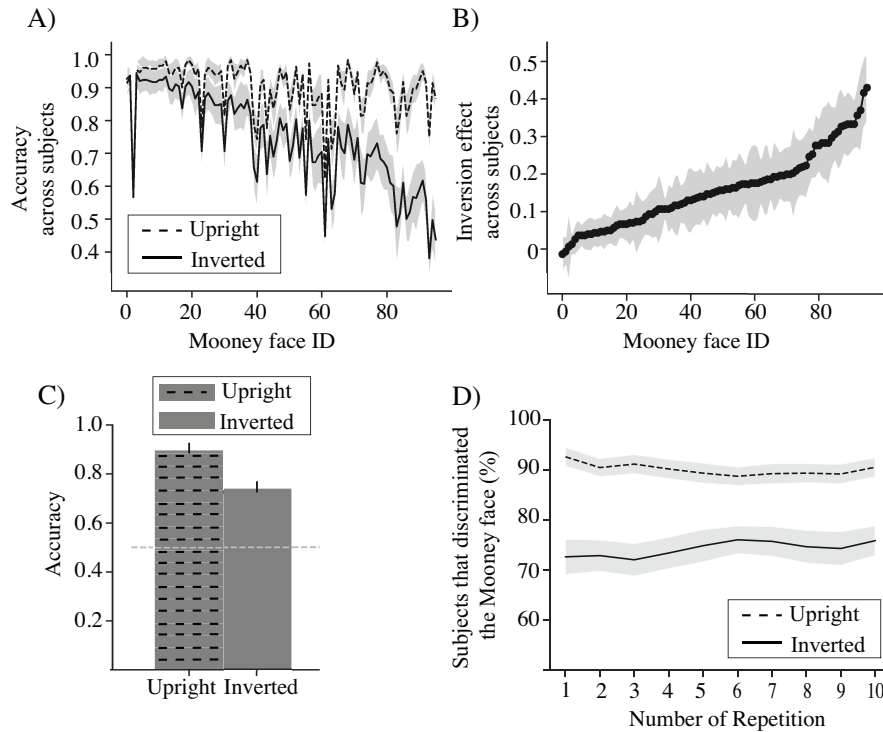


Figure 2.4: Experiment 1 results. (A) Upright (dashed) and inverted (solid) accuracy for each Mooney face. Results for each Mooney face (abscissa) are averaged across subjects. The shaded region represents the 95% CI. (B) Magnitude of the inversion effect (difference between accuracy in upright condition and accuracy in inverted condition from panel A) for each Mooney face (abscissa). The shaded region represents the 95% CI of the mean across subjects. (C) Average accuracy across subjects and across Mooney faces for upright (dashed) and inverted (solid) conditions. The error bars represent the 95% CI of the mean across subjects and Mooney faces. The gray dashed line represents the 50% at chance accuracy. (D) Percentage of subjects that successfully discriminated the Mooney face at each repetition of Mooney face (abscissa). Shaded area represents the 95% CIs of the mean across Mooney faces.

This result suggests that the low between-subject correlation in the inversion effect is driven by low between subject agreement in both the upright and inverted conditions.

To quantify the consistency of these per-face individual differences in holistic processing, we also calculated the within-subject agreement for each Mooney face. The within-subject correlation was performed using a split-half procedure. That is, for each subject and each Mooney face, we separated the upright and inverted trials in two halves (bootstrap with resampling) and then also calculated each face’s inversion effect for each half separately. The within-subject agreement was the correlation between the inversion effect of these two halves. We found that there was a high within-subjects correlation in the inversion effect (Fig 5C; $r = 0.537$, $CI = [0.510, 0.563]$, $p < 0.001$). The high within-subjects correlation shows that subjects were consistent within themselves about the degree to which any particular face was holistically processed (Fig. 5C). The same pattern of results was found for upright and inverted recognition (Figs. 5A-B). This high within-subject consistency reveals that the low between-subject correlation found in Figure 5 was not due to noise.

Lastly, the difference between the within-subjects and between-subjects correlation in the inversion effect was statistically significant ($p < 0.001$, based on bootstrap test). That is, we found consistent per-face individual differences in holistic processing of Mooney faces.

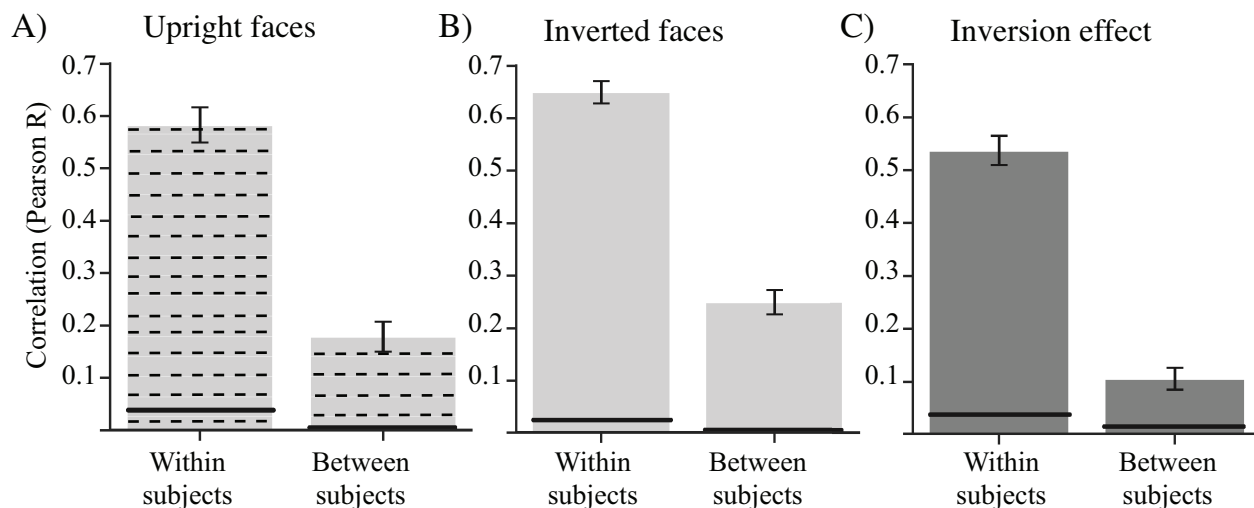


Figure 2.5: Average within- and between-subject agreement in the accuracy scores for the Mooney face task across subjects for the upright condition (A), the inverted condition (B) and the inversion effect (C). The within-subjects agreement is calculated using a bootstrapped split-half correlation (more details in the Results section). The error bars represent the 95% CI of the bootstrapped mean. The black line in each bar represents the upper 97.5% upper bound of the bootstrapped null distribution.

2.3.3 Discussion

Taken together, our results reveal consistent per-subject individual differences in the extent of holistic processing of Mooney faces and stable per-face differences in the extent to which subjects process Mooney faces. Interestingly, we found that subjects do not agree on which Mooney faces are more holistically processed. That is, the Mooney faces that are processed more holistically are not the same for all subjects. These differences are stable within-subject, as reflected by the high within-subject correlations. These idiosyncratic per-face individual differences in holistic processing between subjects do not originate only from individual differences in perceiving the inverted faces.

2.4 Experiment 2: an operational definition of Mooney face ‘holistic-ness’.

Experiment 1 revealed individual subject and per-face differences in holistic processing of Mooney faces. Although observers had stable holistic representations of the Mooney faces within themselves, subjects varied significantly in the extent to which they process Mooney faces holistically. Additionally, Mooney faces showed significantly different levels of holistic processing across subjects. These results suggest that Mooney faces are not all equally good at isolating holistic processing and that some Mooney faces are easier to recognize than others. Experiment 1 showed that there are individual differences in Mooney face perception that operate at the level of particular faces. Since Mooney faces seem to vary in their difficulty and the extent they tap into holistic processing, it is possible that the lack of between-subjects agreement is limited to harder Mooney faces that may require observer-specific high-level representations. To explore this hypothesis, in Experiment 2 we first proposed two ways to identify the harder Mooney faces that required relatively more holistic processing. One way is to use the magnitude of the inversion effect for each Mooney face, which we already calculated in Experiment 1. Another way may be to examine the contours (the outlined edges) of the Mooney faces. Typically, the outlines of Mooney faces are believed to be unrecognizable (Moore & Cavanagh, 1998), but this has not been systematically examined in the literature. Harder Mooney faces that most effectively tap into holistic processing should lack segmentable parts and their contours should therefore be harder to recognize. In contrast, single face feature information should remain in easy, part-based faces, which would still allow them to be recognized. In sum, Experiment 2 aimed to test the hypothesis that the per-face idiosyncratic holistic perception exists for both harder and easier Mooney faces.

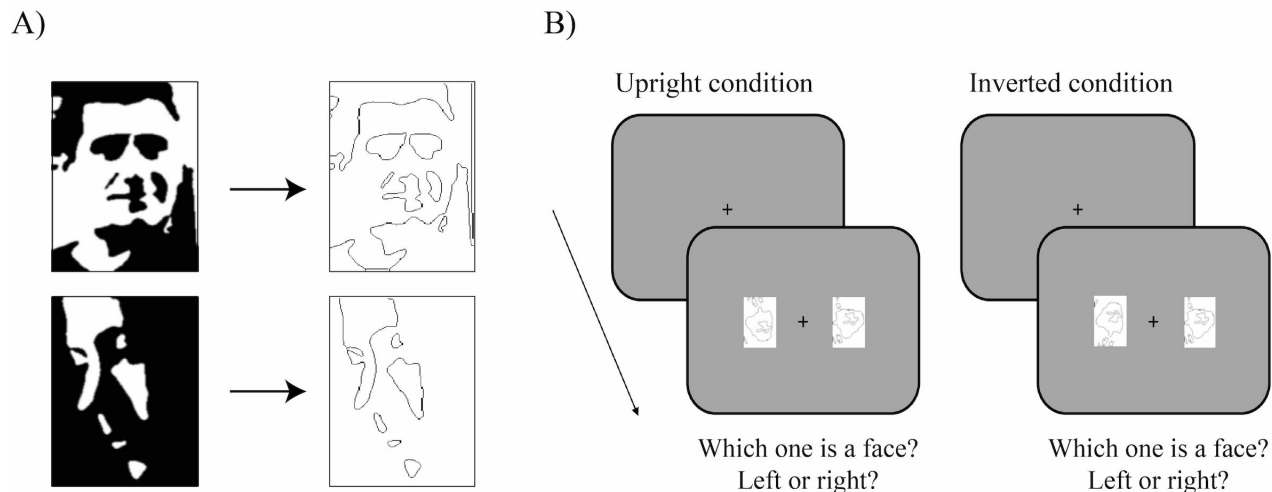


Figure 2.6: Experiment 2 example stimuli and design. (A) Two examples of Mooney faces and their respective outlines. (B) Procedure in the outline condition. The left and right panels represent the upright and inverted conditions, respectively. The full Mooney condition is not depicted. Stimuli, task, and procedure were exactly the same as in Experiment 1.

2.4.1 Method

2.4.1.1 Participants

Twenty-three participants took part in this experiment. All subjects were undergraduate students at the University of California, Berkeley and provided written consent form before participation. All experimental procedures were approved by the UC Berkeley Institutional Review Board.

2.4.1.2 Materials

The Mooney face stimulus is the same as in Experiment 1. An edge-detection algorithm with Sobel-method was run on each of the 192 Mooney face stimuli (scrambled stimuli included). Figure 6A illustrates an example of two Mooney faces and their respective output outline after running the edge-detection algorithm on them.

2.4.1.3 Design

2 Edge conditions (full Mooney vs. outline) x 2 Orientation conditions (upright vs. inverted) x 96 Mooney stimuli x 4 repetitions of each Mooney Face. Subjects ran a total of 1536 trials.

2.4.1.4 Procedure

Task and procedure were the same as in Experiment 1, with the exception that there was an extra condition: the stimuli could be either a Mooney face and a scrambled Mooney face or an outline of a Mooney face and an outline of a scrambled Mooney face. As in Experiment 1, the Mooney face (or the outline of the Mooney face in the outline condition) could be either upright or inverted. Thus, there were four conditions (full versus outlined Mooney face $\times$ 2 orientations). The subject's task was always to determine which of two images was a face regardless of its orientation. The four conditions were randomly interleaved between trials and divided into 10 blocks. Fig. 6B show a sample trial of the outline condition for upright and inverted faces. Trials in the full Mooney face condition were exactly the same as in Experiment 1.

2.4.2 Results

Replicating the inversion effect and the results in Experiment 1, accuracies for the upright condition were overall higher than for the inverted condition in the full Mooney condition, $t(95) = 9.54, p < 0.001$. For both upright and inverted conditions, the accuracy was lower for the outline Mooney than the full Mooney condition, $t(95) = -12.14, p < 0.001$ and $t(95) = -5.40, p < 0.001$.

In Experiment 2 our goal was to investigate whether the per-face individual differences in holistic processing appear in both hard and easy Mooney faces. A hard Mooney face would be one that lacks any low-level, single face features, which means that it would show a strong inversion effect and its outline should be unrecognizable. The per-face inversion effects were pulled from Experiment 1. Here, we used the accuracy in the outlined Mooney condition as a measure of how recognizable the individual features of the face are (eyes, nose, mouth, etc.). The results indicated that some Mooney outlines were easy to recognize, and some were difficult to recognize. Figure 7A shows examples of outlined Mooney faces that were particularly easy to recognize (i.e., showed high accuracy in outlined condition) or were particularly difficult to recognize (i.e., showed low accuracy in outlined condition) with their respective full Mooney versions (Fig. 7B). Then, we correlated the difficulty of each outlined Mooney face with the magnitude of the inversion effect for each respective face, taken from the data Experiment 1 (that is, the two sets of data come from two separate group of subjects).

To test the association between the two measures – the magnitude of the inversion effect and the difficulty of detecting a face in its outline – we pulled the inversion effect values of each Mooney face from Experiment 1 and correlated them with their accuracy in the outlined condition in Experiment 2. That way, each variable came from independent subjects. We found that the variability in recognizability of the outlined Mooney faces was correlated with the degree to which the full Mooney faces required holistic processing. That is, the easier to recognize an outline was (measured by the accuracy of the outline Mooney in upright condition), the weaker the inversion effect for that Mooney face was, Pearson's $r =$

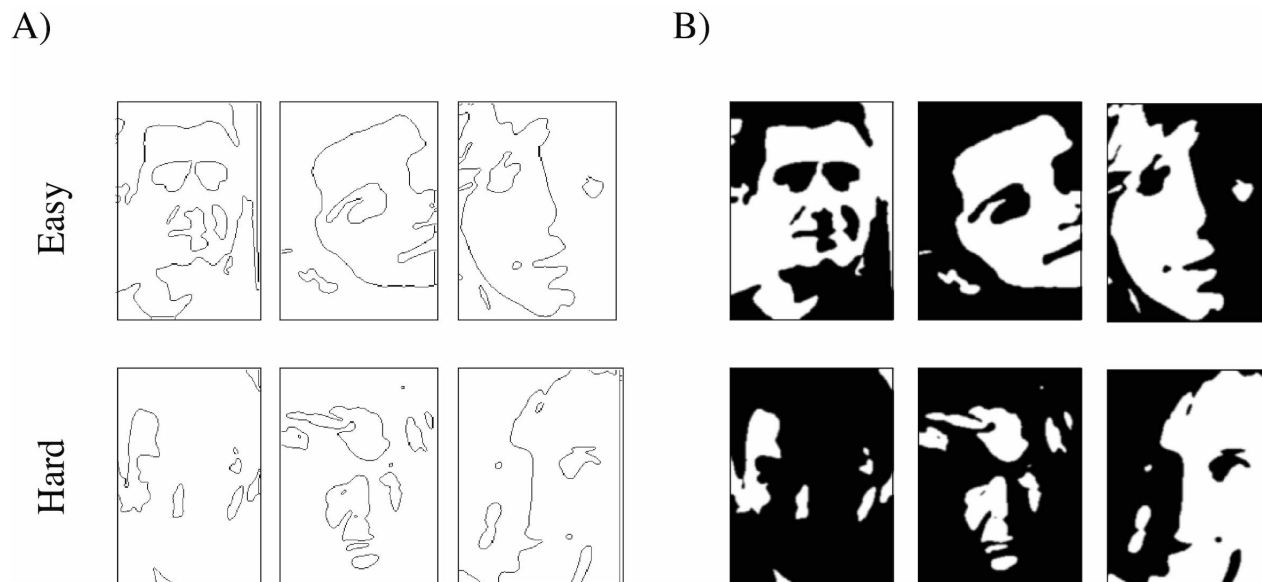


Figure 2.7: (A) Examples of outlines that are easy (upper row; higher accuracy in outlined Mooney condition) and hard (lower row; lower accuracy in outlined Mooney condition) and (B) the respective full Mooney faces.

0.45, $p < 0.001$, $CI = (0.60, 0.28$; Figure 8A). These results suggest that there may be a continuum of how hard a Mooney face is to recognize and how much a Mooney face taps into holistic processing.

Could the per-face individual differences in holistic processing be limited only to hard Mooney faces? In Experiment 1, we found that observers did not agree on the magnitude of the inversion effect for each Mooney face. Here, we wanted to test whether the per-face, per-subject differences found in Experiment 1 are limited to those hard faces that cannot be processed by parts and require holistic processing (i.e., that have a strong inversion effect). Experiment 2 revealed that hard Mooney faces, which tap into holistic processing, can be independently identified by the recognizability of their outlines. Following on this, we correlated the outline recognizability with the level of between subject agreements in the magnitude of the inversion effect of each face. The recognizability of each outlined face was measured by the accuracy in the upright outline condition in Experiment 2. To quantify the between-subjects agreement for each Mooney face, we used the between-subject SD of the inversion effect for each Mooney face in Experiment 1. A SD close to zero represents nearly perfect agreement and the higher the SD gets, the lower the between-subject agreement. Importantly, the two sets of data calculated here – the per-Mooney-face between-subject agreement and the difficulty of each outline face – come from different subjects (from Experiments 1 and 2, respectively). Our results show a negative correlation between the accuracy in recognizing the upright outlined Mooney face and the agreement between subjects, Pear-

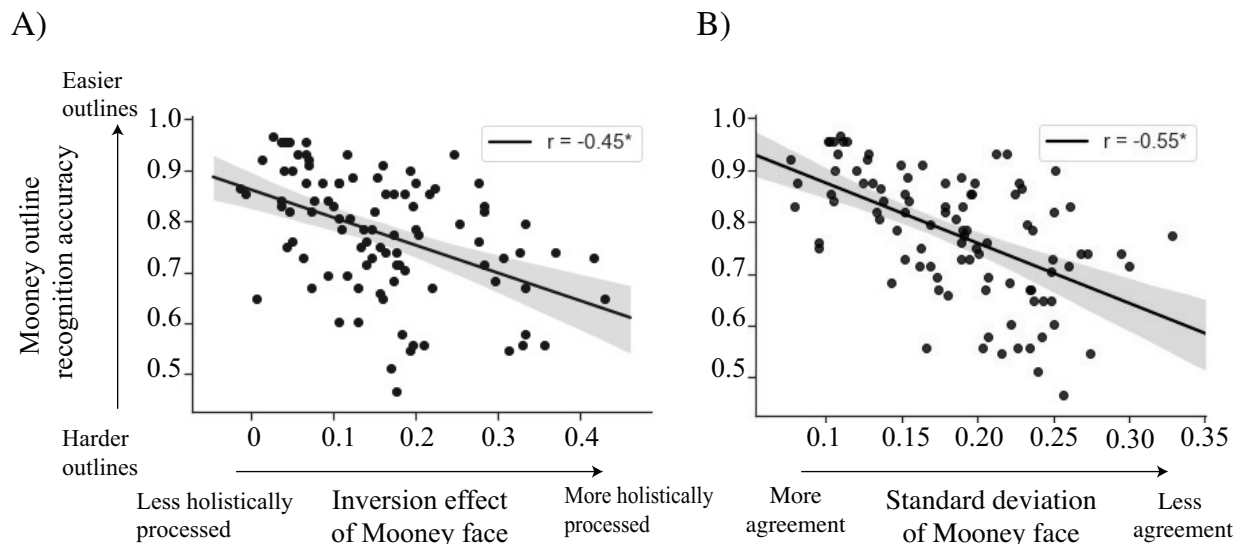


Figure 2.8: (A) Correlation between the magnitude of the inversion effect for each Mooney face from Experiment 1 (abscissa) and the discrimination of the corresponding upright outlined Mooney from Experiment 2. The two axes represent independently collected sets of data from different groups of observers. Each dot represents one Mooney face. The shaded region represents the 95% CIs of a linear regression. There is a correlation between the recognizability of outlines and the holistic-ness of the full Mooney faces. (B) Correlation between the between-subjects agreement in the inversion effect of each Mooney face from Experiment 1 (measured by SD of each Mooney face inversion effect from Experiment 1 – abscissa) and the discrimination of each upright outlined Mooney face from Experiment 2 (ordinate). There is a correlation between the recognizability of outlines and the extent to which subjects agree on the inversion effect of the full Mooney faces. Each dot represents a Mooney face. The shaded region represents the 95% CIs of a linear regression.

son's $r = 0.55$, $p < 0.01$, $CI = (0.67, 0.39)$; Figure 8B). Faces with higher agreement have higher upright accuracy in the outline condition. Faces with less between subject agreements have outlines that are less recognizable. That is, subjects agree more on the easy Mooney faces that had more cartoon-like outlines and weaker inversion effects (Figure 9, top row). In contrast, the individual differences are more apparent in the hard faces that have stronger inversion effects and whose outlines are unrecognizable (Figure 9, bottom row).

2.4.3 Discussion

Our results show that the recognizability of a Mooney face's outline predicts the extent to which that Mooney face is processed holistically. These results confirm and extend the results, we found in Experiment 1 where each Mooney face varied in the extent to which it was



Figure 2.9: Examples of Mooney faces and outlines that differed in inversion effect, outline recognizability, and between-subject agreement. Top row: example Mooney faces with weaker inversion effect, easily recognized outlines, and high between-subject agreement. Bottom row: Mooney faces with stronger inversion effect, difficult-to-recognize outlines, and low between-subject agreement. For both rows, magnitude of the inversion effect and between-subject agreement of Mooney faces are pulled from Experiment 1, while difficulty of discrimination of outlined Mooney faces is taken from Experiment 2.

holistically processed. Those Mooney outlines that are recognizable (e.g., top row, Figure 9) and are essentially seen as cartoon-like. Hard, holistic Mooney faces in the outline condition are literally outlines of blobs. Their full counterpart lacked any single feature information, so their outlines did too. Turning the hard Mooney face into an outline removed the possibility of processing them holistically, which prevents their discrimination. Contrary, outlines of easy, part-based faces maintained the single face feature information, which allowed them to be processed in a part-based manner, and possibly holistically too. Thus, the categorization of how cartoon-like the Mooney face is reveals which faces tap into holistic processing.

Importantly, the results in Experiment 2 suggest that the lack of between-subject agreement in each face's inversion effect found in Experiment 1 is mainly driven by the faces that tap into holistic processing (i.e., have stronger inversion effect) and whose outlines are hard to recognize. In other words, subjects show a higher level of agreement only for those cartoon-like faces that have clearly segmentable features and weaker inversion effects. This echo the results of the first experiment: holistic processing seems to be very subject-specific. Studies aiming to investigate holistic processing using hard Mooney faces that cannot be

processed part-by-part should take into account these individual differences.

2.5 General Discussion

Mooney faces are frequently used to study holistic processing and its role in face recognition (Latinus & Taylor, 2005, 2006; McCaffery et al., 2018; Otsuka, Hill, Kanazawa, Yamaguchi & Spehar, 2012; Verhallen et al., 2014;). The individual differences approach in holistic processing can help elucidate what makes our ability to process faces unique (Yovel, Wilmer & Duchaine, 2014). By examining each observer’s idiosyncratic way of processing faces, we can isolate the face processing mechanisms that are universally shared across humans and those that are idiosyncratic to each observer’s experience. Because Mooney face perception is subject to individual differences (Verhallen et al., 2014), and because these individual differences in holistic processing likely depend on subject-specific top-down knowledge and prior experience (Michel, Rossion, Han, Chung & Caldara, 2006; Zhou, Cheng, Zhang & Wong, 2012), one might expect that there should be both individual subject-specific differences as well as individual stimulus-specific differences in holistic processing. Consequently, some particular Mooney faces may be easier to recognize for some particular subjects. To address this, we tested whether there are individual differences in the extent to which humans process Mooney faces holistically and whether Mooney faces vary in the extent to which they tap into holistic processing by calculating the magnitude of the inversion effect of each individual Mooney face.

Our results showed that there are individual differences in the extent to which subjects process Mooney faces holistically. This finding replicates previous studies that include not only Mooney faces as a measure of holistic processing, but also use the inversion effect in gray-scale faces (Verhallen et al., 2014; McCaffery et al., 2018), the composite-face effect (Richler, Cheung & Gauthier, 2011) and the whole-part effect (Wang et al., 2012). Importantly, observers did not agree on which faces had a weak or strong inversion effect; that is, which faces were holistically processed. Yet, observers did show within-subject consistency in their holistic perception of the Mooney faces. These stable per-face idiosyncratic differences indicate that the holistic representations of Mooney faces are not universal or shared across subjects.

Furthermore, we found that upright Mooney faces were either recognized in the first repetition or not recognized at all, and once an observer recognized a Mooney face they could not “unsee” it (Brodski et al., 2015), much like Dallenbach’s cow (Dallenbach, 1951) and Gregory’s Dalmatian dog (Gregory, 1970). Note that inverted faces did show an improvement with repeated face exposures. Here, we extend these results to Mooney faces and show that the high within-observer consistency occurs for specific faces. Mooney faces, unlike other kind of faces, may require high-level representations that are observer-specific. If an observer has a template that facilitates the discrimination of a particular Mooney face, then the observer will recognize it in the first trial; otherwise, the face will not be recognized regardless of the number of repetitions. These face templates may be shaped by individual experience, which

would make them unique to each observer.

Previous studies suggest that faces and objects lie on a continuum, such that faces are processed as a whole and objects are processed by their parts (Farah, 1991). The results from Experiment 1 showed that Mooney faces vary significantly in how much they tap into holistic processing. That is, not all Mooney faces are processed holistically to the same extent. Our results suggest that given a Mooney face stimulus set, one can filter out which faces can be recognized using feature-based processing by independently measuring the recognizability of the outlines. Several research groups have created Mooney faced datasets by either directly drawing them (Mooney, 1957), thresholding grayscale faces (Verhallen & Mollon, 2016; Schwiedrzik et al., 2018) or generating them through machine learning algorithms (Ke et al., 2017). However, none of them include an evaluation of the “holistic-ness” of the Mooney faces.

The approach developed here shows that one can use the contour of the face to predict how holistic it is on a group or even individual subject basis. Although Mooney contours are often considered ambiguous (Moore & Cavanagh, 1998), our results show that some Mooney faces do have recognizable contours. Many Mooney outlines are completely unrecognizable, but some appear cartoon-like. The cartoon-like Mooney outlines were associated with easily recognized Mooney images. Conversely, the outlines that were unrecognizable were derived from Mooney faces that required more holistic processing. Using the technique here and measuring outline recognizability can help future studies isolate holistic processing by selecting the most appropriate Mooney faces.

Importantly, quantifying the “holistic-ness” of the Mooney faces using this approach revealed that per-face, per-observer differences in holistic processing only occurred for hard Mooney faces that required holistic processing. Conversely, discrimination of easily segmentable Mooney faces was shared across subjects and seems to enjoy more universal agreement. These findings suggest that holistic processing of hard faces depends on a particular observer’s experience or idiosyncratic templates, whereas processing of easy faces proceeds similarly across observers. Top-down guidance, upon which holistic face perception of hard faces may rely, depends on knowledge and prior experience. Since knowledge and prior experience is unique to the observer, the individual differences are more apparent in those faces that are harder to recognize and require more holistic processing. Altogether, these results support the conclusion that holistic and part-based processing are distinct mechanisms of face recognition (Moscovitch, Winocur Behrmann, 1997; McKone, 2004). Here, we provide further evidence for a dissociation between these two mechanisms: part-based processing is universal, while holistic perception is individual-specific and depends on experience.

A large number of behavioral and neuroimaging studies have aimed to understand holistic perception and its role in face recognition. Many of these studies use Mooney faces and presume that holistic processing is involved (Kanwisher et al., 1998; Andrews & Schluppeck, 2004; George et al., 2005; Verhallen et al., 2014; Verhallen & Mollon, 2016; McCaffery et al., 2018). Some of this previous work has taken into consideration the idea of individual differences in the extent to which Mooney faces are recognized (Andrews & Schluppeck, 2004; George et al., 2005), in agreement with our results. However, Mooney face awareness

was collapsed across observers in these studies, which did not allow for a further analysis of the stability of these differences or the possible stimulus-specific individual differences. Our results show that the degree to which a Mooney face is processed holistically is idiosyncratic: it depends on the particular face and the particular observer. Therefore, to really understand the nature of holistic representations, future research should take this idiosyncrasy into account.

2.6 Acknowledgments

The authors are grateful for helpful discussions with Ken Nakayama, William Prinzmetal and Ervin Hafter. The authors also want to thank Aneesa Conine-Nakano for her assistance in data collection.

Chapter 3

Relative Tuning of Holistic Face Processing Towards the Fovea

3.1 Abstract

Humans quickly detect and gaze at faces in the world, which reflects their importance in cognition and may lead to tuning of face recognition toward the central visual field. Although sometimes reported, foveal selectivity in face processing is debated: brain imaging studies have found evidence for a central field bias specific to faces, but behavioral studies have found little foveal selectivity in face recognition. These conflicting results are difficult to reconcile, but they could arise from stimulus-specific differences. Recent studies, for example, suggest that individual faces vary in the degree to which they require holistic processing. Holistic processing is the perception of faces as a whole rather than as a set of separate features. We hypothesized that the dissociation between behavioral and neuroimaging studies arises because of this stimulus-specific dependence on holistic processing. Specifically, the central bias found in neuroimaging studies may be specific to holistic processing. Here, we tested whether the eccentricity-dependence of face perception is determined by the degree to which faces require holistic processing. We first measured the holistic-ness of individual Mooney faces (two-tone shadow images readily perceived as faces). In a group of independent observers, we then used a gender discrimination task to measure recognition of these Mooney faces as a function of their eccentricity. Face gender was recognized across the visual field, even at substantial eccentricities, replicating prior work. Importantly, however, holistic face gender recognition was relatively tuned—slightly, but reliably stronger in the central visual field. Our results may reconcile the debate on the eccentricity-dependence of face perception and reveal a spatial inhomogeneity specifically in the holistic representations of faces.

3.2 Introduction

Faces play a central role in our everyday life. Face perception is at the center of most social interactions, from allowing us to recognize our friends among a group of strangers to facilitating communication through non-verbal language and emotion recognition (Jack & Schyns, 2015). Humans tend to quickly direct their attention and gaze to the faces in a scene (Bindemann et al., 2005; Cerf et al., 2009; Cerf et al., 2008; Costela & Woods, 2019; Coutrot & Guyader, 2014; Foulsham et al., 2010; Jack & Schyns, 2015; Marat et al., 2013; Ro et al., 2001; Theeuwes & Van der Stigchel, 2006) and we prefer faces to other objects from a very young age (Umiltà et al., 1996; Farzin et al., 2009; Frank et al., 2009). Indeed, we are highly trained with faces, as we typically see and recognize many faces every day starting in infancy. Face perception therefore enjoys a unique and privileged role in cognition.

One of the main advantages faces have over other objects is that we process them holistically – that is, we process faces as a whole instead of (or in addition to) processing the segmentable parts that make the face (eyes, nose, mouth, etc; Farah et al., 1998; Sergent, 1984). Holistic processing develops early in humans and reaches adult-like levels in childhood (Crookes & McKone, 2009). Holistic processing has been traditionally operationalized by the magnitude of the inversion effect, the enhanced recognizability of upright faces relative to inverted ones. While upright faces are processed holistically, inverted faces are perceived via part-based processes (Yin, 1969). This gives upright faces an advantage over inverted faces—they are processed faster and more accurately (Maurer et al., 2002; Rossion, 2008; Rossion, 2009).

Because of the critical role they play in our lives, faces are quickly identified in a scene and are often processed at the fovea. The resulting life-long experiences we have foveating faces, and the perceptual learning that accompanies this (Bi, Chen, Weng, He & Fang, 2010; Gold, Bennett & Sekuler, 1999), might therefore produce specialized mechanisms that are uniquely found in the central visual field. That is, face recognition may be retinotopically tuned to the fovea. Unfortunately, previous work has yielded mixed results on this question. Behavioral and psychophysical results have found little foveal selectivity in face recognition, whereas neuroimaging studies have provided evidence for a central-field bias in face representations.

On the one hand, previous behavioral work has shown that face recognition is relatively unaffected in the periphery up to 16° of eccentricity (McKone, 2004), at least when there is little clutter or crowding (Louie, Bressler & Whitney, 2007). Extrafoveal faces can be detected, recognized, selected, and saccades can be made to their peripheral locations faster and more accurately than to other objects (Boucart, Lenoble, Quetelart, Szaffarczyk, Despretz & Thorpe, 2016; Crouzet, Kirchner & Thorpe, 2010; Hershler, Golan, Bentin & Hochstein, 2010; Wolfe & Whitney, 2014). In line with these findings, holistic processing of faces also happens across the visual field (Kovács, Knakker, Hermann, Kovács & Vidnyánszky, 2017; Kreichman, Bonnef, & Gilaie-Dotan, 2020; McKone, 2004). To investigate the extent to which faces are processed holistically in the periphery, McKone (2004) compared recognition of a Mooney face and a single nose in five visual field locations. McKone (2004) found that holistic processing happened at all tested eccentricities (up to 21 degrees in left and right

visual fields) only for the whole faces. Processing of a single face feature (the nose) yielded little to no holistic processing at any eccentricity. Similar conclusions have been reached by others (Kovács et al., 2017; Kreichman et al., 2020).

On the other hand, neuroimaging work has revealed a central-field bias in face representations that does not exist as strongly for other objects (Baldassano, Iordan, Beck & Fei-Fei, 2012; Hasson, Levy, Behrmann, Hendler & Malach, 2002; Kanwisher, 2001; Kamps, Hendrix, Brennan & Dilks, 2020; Levy, Hasson, Avidan, Hendler & Malach, 2001; Sayres & Grill-Spector, 2008; Silson, Groen, Kravitz & Baker, 2016). For example, Schwarzlose and colleagues found that there was stronger BOLD response for foveally presented faces (Schwarzlose, Swisher, Dang & Kanwisher, 2008), and Nichols and colleagues found that face decoding was better in the central compared to peripheral positions (Nichols, Betts & Wilson, 2010). Similarly, population receptive fields for face stimuli in several occipito-temporal areas are biased toward central and lower visual field locations (Gomez, Gomez, Natu, Jeska, Barnett & Grill-Spector, 2018; Silson et al., 2016; Kay, Weiner & Grill-Spector, 2015; Yue, Cassidy, Devaney, Holt & Tootell, 2011). Some of the foveal bias in face representations might arise from factors like cortical magnification and receptive field scaling, or to confounds including SNR differences within and across brain regions. However, face and object selective regions are highly selective to visual location, even in the periphery (e.g., multivoxel pattern analysis revealed retinal selectivity of 0.14E in FFA, which indicates that a face shifted by 14% of its eccentricity was reliably discriminated, even for objects at 9 degrees eccentricity; Fischer & Whitney, 2011). In addition to this, other confounds like variations in SNR are not thought to be entirely responsible for the visual field biases in face representations (Silson et al., 2016, Poltoratski, Kay & Grill-Spector, 2019). There is therefore sufficient and converging evidence to suggest that face representations are stronger or biased (or perhaps remapped; Williams et al., 2008; Edwards, VanRullen & Cavanagh, 2018; Wolfe & Whitney, 2014) toward the fovea (Hasson et al., 2002, Kay et al., 2015, Yue et al., 2011, Silson et al., 2016, Sayres & Grill-Spector, 2008). Such biases in face representations may speak to the importance of perceptual learning: we tend to foveate faces in natural scenes, as they have crucial information for social communication and interaction (Costela & Woods, 2019, Jack & Schyns, 2015, Coutrot & Guyader, 2014). However, the neuroimaging findings beg the question of whether there is some behavioral consequence of the central visual field bias in face representations, and why previous psychophysical results have not clearly found it.

Here, we propose addressing these mixed results by testing for any degree of foveal tuning of holistic face perception. In other words, we hypothesize that there may be a relative preference or relative strength of holistic processing of faces at the fovea compared to the periphery. To isolate holistic processing, we used a more stringent stimulus-specific operational definition of holistic perception than is typically used. To test for the eccentricity-dependence of face gender recognition, we controlled for the degree of holistic processing required by each specific stimulus. Previous work shows that some faces tap into holistic perception more readily than others (Canas-Bajo & Whitney, 2020), so we isolated the most holistically processed faces. In the following experiment, we used Mooney faces (Fig. 1A), which are

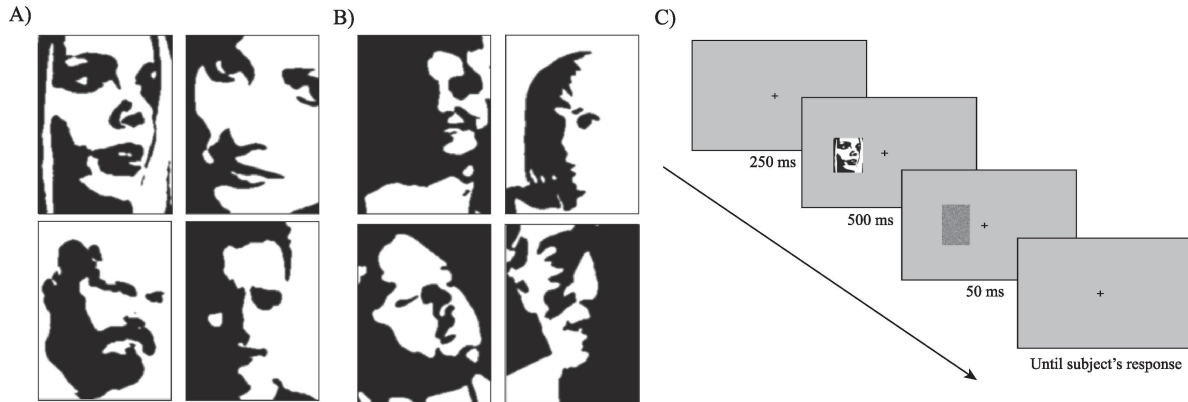


Figure 3.1: A) Four sample two-tone faces in the low-holistic group. B) Four sample Mooney faces in the high-holistic group. The two groups were formed based on the results of pilot study B with independent subjects (details in the Methods section). All stimuli were extracted from Schwiedrzik et al. (2018). C) Illustration of experiment procedure. Faces not drawn to scale.

two-tone images that are readily perceived as faces even though they lack segmentable face features (Mooney, 1957). Mooney faces are a special set of two-tone images because they maintain the minimal stable structure of faces (Ke et al., 2017). Because these faces don't have easily segmented face features, they have to be processed as a whole to be perceived as faces, which makes them ideal stimuli to test holistic perception (Andrews & Schluppeck, 2004; Bona, Cattaneo Silvano, 2016; Cavanagh, 1991; Farzin et al., 2009; Latinus & Taylor, 2005; McKone, 2004; Moore & Cavanagh, 1998). However, previous work shows that there are some two-tone images that do retain some featural information and can be processed by parts and/or holistically in parallel (Canas-Bajo & Whitney, 2020). These two-tone faces would therefore not be ideal Mooney faces, per se, since they do not isolate holistic processing. In contrast, purely holistic Mooney faces can only be processed holistically: one must first recognize the whole image as a face before any particular feature or fragment can be identified as, for example, an eye, nose, or mouth (Cavanagh, 1991). Likewise, one must also recognize the image as a face before the gender or identity of the Mooney face can be discriminated (Cavanagh, 1991; Farzin et al., 2009). By testing the extent of holistic processing across the visual field for relatively holistic faces, we aim to investigate 1) whether there is indeed a central-field bias for face processing and 2) whether it exists exclusively for holistic faces. To foreshadow our results, faces are readily recognized throughout the visual field, and, interestingly, we found a slight foveal tuning for purely holistic Mooney face gender recognition, such that holistic Mooney face perception was relatively stronger near the fovea.

3.3 Method

3.3.1 Participants

Twenty-four participants took part in the experiment. All subjects were undergraduate students at the University of California, Berkeley and provided written consent before the start of the experiment. The UC Berkeley Institutional Review Board granted ethics approval for the study.

3.3.2 Material

The stimuli consisted of forty images of Mooney faces extracted from Schwiedrzik et al., (2018). To select these faces, we ran two pilot experiments (pilot experiment A and B) with independent observers: one that tested the gender of the faces and another one that measured the extent to which the faces tapped into holistic processing. The presentation of the stimuli and the data collection of both pilot experiments was done online and was controlled using Qualtrics (<https://www.qualtrics.com>). All stimuli presented in the pilot experiments were of size 160×230 pixels but note that participants used different monitors and distance to monitor could not be controlled.

In pilot experiment A, our goal was to determine the ground truth gender of each face. Because we did not have access to the original gray-scale face images from which the Mooney faces were generated, and because gender appearance is subjective in any case (Freeman, Rule, Adams & Ambady, 2010), we asked twenty independent observers to categorize the gender of a random sample of 274 faces from the original Schwiedrzik et al. (2018) Mooney dataset (about half of that original dataset). Each trial consisted of a Mooney face that was presented at the center of the screen until participants recorded a response. The gender of 216 faces was rated with 90

In pilot experiment B, our goal was to measure how holistic each Mooney face image was. The present study aimed to compare eccentricity-dependent performance in two groups of faces: high-holistic faces and low-holistic faces. High holistic Mooney faces are those that isolate holistic processing while low holistic two-tone faces are those that can be processed holistically and in a part-based manner. Following the procedure in Canas-Bajo & Whitney (2020), to quantify the extent to which each face tapped into holistic processing, we measured the magnitude of the inversion effect for each face in pilot experiment B with twenty new, independent observers. Here, participants were asked to categorize the gender of the 216 Mooney faces selected in pilot experiment A. Each face was presented once upright and once inverted across the experiment, and the inversion effect for each face was quantified as the difference in accuracy between the upright and inverted condition. Faces were presented foveally until participants recorded a response. Participants were allowed to move their eyes. Overall, we found the well-known inversion effect: gender recognition of upright faces was more accurate (Mean: 0.97, SEM: 0.0) than of inverted faces (Mean: 0.85, SEM: 0.01). The magnitude of the inversion effect however varied across Mooney faces, replicating Canas-Bajo

& Whitney (2020). Out of the 216 faces tested, we selected the 20 faces with the weakest inversion effect (Mean: 0.0, SEM: 0.0) and the 20 faces with the strongest inversion effect (Mean: 0.41, SEM: 0.03), $t(19) = 0.0$, $p < 0.001$. From now on, we will refer to these two groups of faces as the “low-holistic” group (Fig. 1A) and “high-holistic” group (Fig. 1B), respectively.

In the main study, stimuli were presented on a CRT monitor at 100 Hz refresh rate, with 1024×768 pixels resolution and a horizontal screen size of 40.5 cm. The monitor was placed 60 cm from a chin rest that stabilized the participant’s head. At this distance, all the face stimuli shown during the experiment subtended 6 degrees visual angle. The presentation of the stimuli was controlled using MATLAB R2016b with Psychophysics Toolbox 3 (Brainard, 1997, Kleiner et al., 2007).

3.3.3 Design

Each subject completed a total of 1300 trials: (2 orientation conditions (Upright vs. Inverted) $\times$ 2 face conditions (High-holistic vs. low-holistic face group) $\times$ 13 locations (2° , 4° , 6° , 8° , 10° and 12° in both the right and left visual field, as well as the fovea) $\times$ 25 repetitions. Each orientation-face-location condition was repeated 25 times.

3.3.4 Procedure

Participants performed a gender-discrimination task in which they had to determine the gender of an upright or inverted Mooney face at different visual field locations. We chose a gender discrimination task because it was ideal for our goals and stimuli: First, we needed unfamiliar faces and yet Mooney faces, like shadow-defined shapes, are not identifiable unless they are very familiar identities (Moore & Cavanagh, 1998). Second, a face detection or face/non-face discrimination task would have required relying on “scrambled” face lures—tasks and stimuli that are subject to, and can introduce, bias. Lastly, the goals of our study necessitated a task that required holistic processing, and Mooney gender discrimination requires holistic processing (Sergent & Signoret, 1992; Steeves et al., 2006; Yokoyama, Noguchi, Tachibana, Mukaida & Kita, 2014; Zhao & Hayward, 2010). Although previous literature using grayscale faces shows that featural information may be used to determine the gender of a face (Brown & Perrett, 1993; Schyns, Bonnar & Gosselin, 2002; Nestor & Tarr, 2008; Hu, Hu, Xu & Qin, 2013; Yamaguchi, Hirukawa & Kanazawa, 2013), others have shown the importance of holistic processing in face gender discrimination (Yokoyama et al., 2014, Zhao and Hayward, 2010). More importantly, because we used Mooney faces, holistic processing plays a key role in our face gender discrimination task (Sergent & Signoret, 1992, Steeves et al., 2006, Yokoyama et al., 2014, Zhao & Hayward, 2010).

A trial sequence started with a 250 ms (ms) fixation cross. Next, a face was shown for 500 ms at one of the thirteen possible locations. Participants were instructed to ignore the orientation of the face. Immediately after the face disappeared, a noise mask was shown for 50 ms. A grey blank screen was displayed until a response was given (self-paced). The trial

ended with a 500 ms inter-trial-interval. Subjects were asked to fixate at the center of the screen throughout the trial and avoid any eye movements. Fig. 1C illustrates a summary of the design and procedure of the experiment.

3.3.5 Data analysis

Here, we operationalized the extent of holistic processing by the magnitude of the inversion effect (Yin, 1969). The inversion effect for each location and face group (high vs. low-holistic) was calculated by subtracting the mean accuracy in the inverted from the upright conditions. We performed a three-way repeated measures ANOVAs with Orientation (upright vs. inverted), Face group (high vs. low-holistic group) and Eccentricity (13 locations tested) as factors and accuracy as the dependent variable. We also performed a two-way repeated measures ANOVA with Face group and Eccentricity as factors and magnitude of the inversion effect as the dependent variable. To test the significance of some effects, we calculated the null distributions using permuting and bootstrapping procedures with 10,000 iterations (Efron & Tibshirani, 1994). All reported pair-comparisons p values were corrected using Benjamini–Hochberg (BH) procedure (Benjamini & Hochberg, 1995).

3.3.6 Results

The main goal of this experiment was to investigate the eccentricity-dependence of face gender recognition and whether it depends on the “holisticness” of each face. Here, we operationalized holistic processing as the magnitude of the inversion effect. The inversion effect is the accuracy in the upright condition minus the accuracy in the inverted condition. Thus, we first calculated the average in the upright and inverted conditions, separately, for each face group at each location. We applied a three-way (Orientation $\times$ Face group $\times$ Eccentricity) repeated measures ANOVA for upright and inverted accuracies.

Both low- and high-holistic upright faces were recognized above chance in the periphery (maximum tested eccentricity was 12 degrees; Low holistic: mean = 0.89, SEM = 0.01; High holistic: mean = 0.70, SEM = 0.02; Fig. 2A-B). This result replicates and extends prior work (e.g., McKone, 2004). For both low- and high-holistic groups, the accuracies for the upright condition were higher than for the inverted condition, $t(23) = 12.87$, $p < 0.001$, $d = 0.24$ and $t(23) = 25.06$, $p < 0.001$, $d = 0.69$, replicating the well-studied inversion effect. Unsurprisingly, the difference between upright and inverted faces at each location was larger in the high-holistic group than in the low-holistic group, revealed by a significant interaction between Orientation and Face group, $F_{2,46} = 174.50$, $p < 0.001$, $d = 0.08$. This was expected, because the division between high- and low-holistic face stimuli was based on independent observers in pilot testing (see Methods). In fact, gender recognition of high-holistic inverted faces remained close to chance across all eccentricities (Mean = 0.56, SEM = 0.02; Fig. 2A); this conforms to the definition of a holistic face being one that is easily recognizable when upright but not when inverted. In contrast, low-holistic faces were recognized well in the periphery with almost no decrement up to 12 degrees in both upright and inverted

conditions (Upright: Mean = 0.89, SEM = 0.02; Inverted: Mean = 0.81, SEM = 0.02; Fig. 2B). Importantly, we found a significant interaction between face group and eccentricity for the upright condition ($F_{12,276} = 3.48$, $p < 0.001$, $\eta^2 = 0.03$). Accuracy for upright high-holistic faces was 0.85 (SEM = 0.02) at the fovea, but then performance decreased as faces appeared further in the periphery (Fig. 2A, red curve). In contrast, gender recognition of upright low-holistic faces remained relatively unaffected in the periphery compared to the fovea (Upright: Mean = 0.89, SEM = 0.02; Fig. 2B, red curve). This result suggests a relative foveal tuning in holistic processing of upright faces: a slightly stronger preference for purely holistic faces at the fovea, compared to the periphery.

To clarify the interaction between eccentricity and degree of holistic processing, we quantified the magnitude of the inversion effect for each face group at each eccentricity. A two-way repeated measures ANOVA on the magnitude of the inversion effect revealed a significant effect of Face group, $F_{1,23} = 70.32$, $p < 0.001$, $\eta^2 = 0.26$. Unsurprisingly, we found that high-holistic faces had an overall stronger inversion effect than low-holistic faces, $t(24) = 7.74$, $p < 0.001$, $\eta^2 = 0.53$. As discussed above, the a-priori hypothesis is that there may be an effect of eccentricity on face gender recognition. There is no expectation of a left versus right visual field difference, and, indeed, there is no significant difference between left and right visual fields for any condition (all p s > 0.05). Because our primary interest is eccentricity, we collapsed across isoeccentric locations in the left and right visual field (Fig. 2D). Low-holistic faces remained relatively unaffected up to 12 degrees of eccentricity while high-holistic faces showed a relatively stronger inversion effect at the fovea that decreased gradually with eccentricity. There was a significant interaction between Face group and Eccentricity (seven eccentricities total), $F_{6,138} = 2.97$, $p < 0.01$, $\eta^2 = 0.03$.

To help visualize this interaction, we calculated the holistic difference at each eccentricity (Fig. 2E). That is, the difference in magnitude of the inversion effect between high and low-holistic faces at each eccentricity. The holistic difference revealed that there was a larger difference between high and low-holistic face groups at the fovea than other eccentricities (Fig. 2E). To further test the significance of this interaction, we calculated the null distribution for the interaction by shuffling the face group label for each location in a nonparametric permutation procedure. We found that the difference in inversion effect between high and low-holistic face group was significant at all locations up to 10 degrees ($p < 0.05$ per permutation analysis) and importantly, was largest at the fovea (compared to 6, 10 and 12 degrees, $p < 0.05$ BH-corrected; Fig. 2E). A linear regression fit to the data further confirmed this result ($n = 168$, $R^2 = 0.07$, $p < 0.01$). This result further suggested that there is some degree of foveal tuning of face gender recognition that is specific to holistic faces.

In addition, we tested whether the foveal tuning found in high holistic faces was due to floor performance in high holistic inverted faces across the visual field. The goal of this analysis was to rule out the possibility that floor or ceiling performance alone may cause the relatively stronger foveal tuning in high holistic faces. To this end, we identified inverted faces with an accuracy above 50% (average accuracy 60% across all eccentricities; all p s > 0.05 ; Fig. 3A). These selected high and low holistic faces are not at floor. Next, we calculated the inversion effect for the low and high holistic groups separately including only those faces.

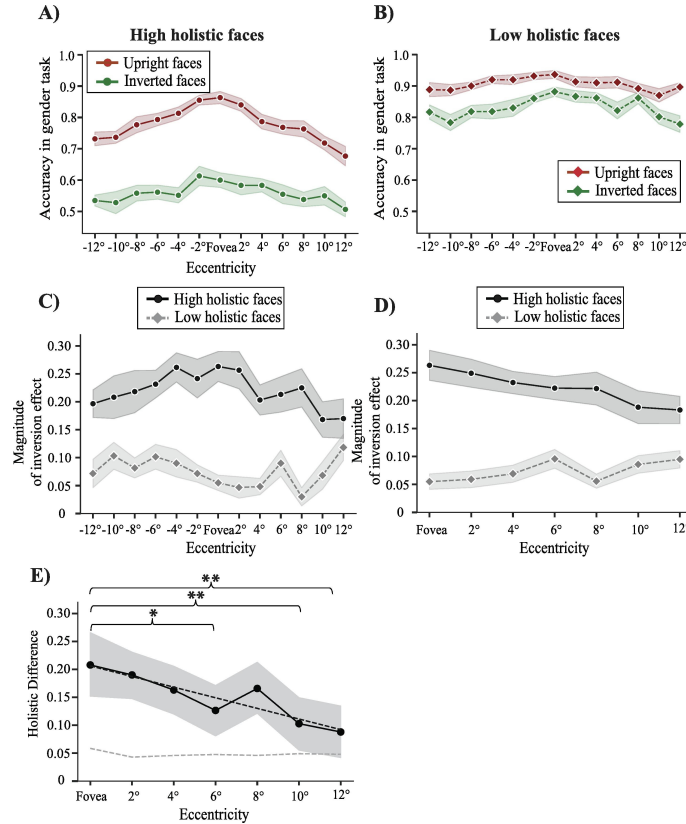


Figure 3.2: Results of Experiment 1. A) Gender recognition of high-holistic faces that were upright (red) or inverted (green). Eccentricity is shown on the abscissa. B) Gender recognition of low-holistic faces that were upright (red) or inverted (green) as a function of eccentricity. C) Magnitude of the inversion effect for high- (dark grey) and low- (light grey) holistic faces. The inversion effect is operationalized as the difference between accuracy for upright minus inverted faces at each eccentricity. D) Magnitude of the inversion effect as a function of unsigned eccentricity for high- (dark grey) and low- (light grey) holistic faces. The shaded region represents the standard error of the mean between subjects in panels A-C. E) Difference between inversion effect (blue) for high versus low holistic faces. Stars represent the significance of the pair comparisons of holistic difference between eccentricities, one star represents $p < 0.05$, two stars represents $p < 0.01$, BH-corrected. The shaded grey region represents the bootstrapped 95% confidence intervals. The dashed-dotted black line represents the linear regression fit to the data. The dashed grey line represents the 97.5% upper bound of the permuted null difference.

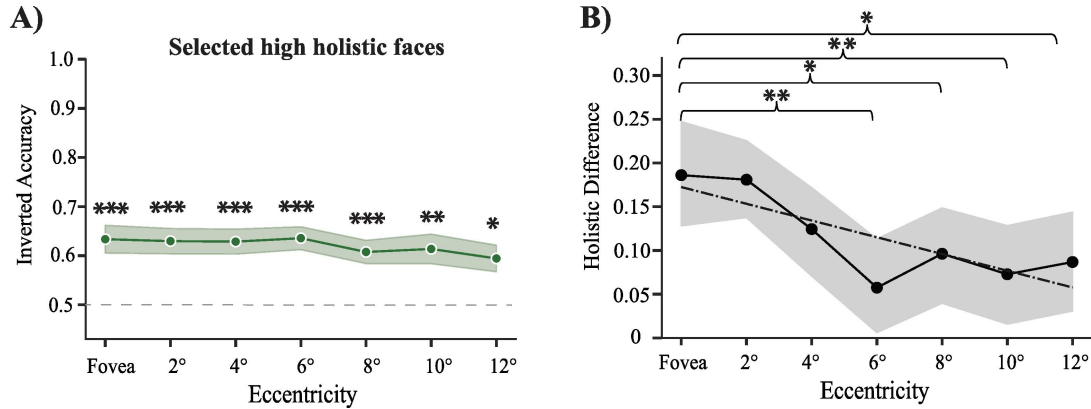


Figure 3.3: Control analysis. A) Inverted faces that are recognized greater than 50% of the time were selected for subsequent analyses. The shaded region represents the standard error of the mean between subjects. Asterisks indicate statistical significance (one star represents $p < 0.05$, two stars represents $p < 0.01$, three stars represent $p < 0.001$, BH-corrected). B) Difference between inversion effect for high versus low holistic faces that were selected. The ordinate shows the holistic difference score (as in Fig. 2E). Stars represent the significance of the paired comparisons of holistic difference between the fovea and each eccentricity, one star represents $p < 0.05$, two stars represents $p < 0.01$, BH-corrected. The dashed-dotted black line represents the linear regression fit to the data. The shaded grey region represents the bootstrapped 95% confidence intervals.

Lastly, we tested whether this group of above-floor faces showed similar foveal tuning specific to high holistic faces by examining the difference in magnitude of the inversion effect between high and low-holistic faces at each eccentricity. If the at-chance inverted accuracy in high holistic faces drives the pattern of foveal tuning, then we should see that this pattern does not persist when we limit the analysis to faces that have significantly above-chance accuracy. The holistic difference revealed that high holistic faces were recognized more readily at the fovea than at any other eccentricity, and this persisted when we restricted the analysis to faces with above-chance inverted accuracy (Fig. 3B). There was a significant difference for fovea versus 6-, 8-, 10-, and 12-degree eccentric faces (all $ps \leq 0.05$, BH-corrected). A linear regression fit to the data confirm this tuning ($n = 168$, $R^2 = 0.06$, $p < 0.01$). In sum, our results show that there is a relative foveal tuning in holistic face processing of upright faces that is not explained by a floor performance in inverted faces.

3.4 Discussion

In the present study, our goal was to reconcile an existing debate on the eccentricity-dependence of face perception. Behavioral studies have shown that faces are relatively un-

affected by the constraints of peripheral vision and that holistic processing happens across the visual field (Kovács et al., 2017; Kreichman et al., 2020; McKone, 2004) whereas brain imaging studies have provided evidence for a central-field bias in faces that does not exist for other objects (Amedi, Malach, Hendler, Peled Zohary, 2001; Gomez et al., 2018; Grill-Spector & Weiner, 2014; Hasson et al., 2002; Kanwisher, 2001; Kay et al., 2015; Levy et al., 2001; Malach, Levy Hasson, 2002; Silson et al., 2016; Yue et al., 2011). Here, we aimed to resolve this controversy by investigating to what extent the eccentricity-dependence of face perception is determined by the degree to which holistic processing is required to process each face. In order to isolate holistic processing from faces that can be processed by parts, we categorized two-tone faces into high-holistic Mooney faces (that required holistic processing) and low-holistic two-tone faces (that relied less on holistic processing and can be processed by parts and/or holistically), following the method in Canas-Bajo & Whitney, 2020. Replicating previous work in the field, we found that upright faces, holistic or not, are well recognized at 12 degrees in the periphery or beyond (Farzin et al., 2009; Kovács et al., 2017; McKone, 2004). Also, in accordance with previous findings, our results show that holistic processing occurs for peripheral faces, even at very large eccentricities (McKone, 2004). The novel result in our study is the dissociation in the effect of eccentricity on high and low-holistic face gender recognition. We observed that holistic faces were foveally tuned, which means that holistic processing was slightly but reliably stronger at the fovea than in the periphery. Together, our results shed light on the discrepancy between psychophysical and brain imaging studies: there is indeed some degree of face processing tuned toward the fovea, but it is particularly stronger for purely holistic faces.

In natural visual search and recognition tasks, observers tend to prefer to foveate faces (e.g., Boucart et al., 2016; Martin, Davis, Riesenhuber & Thorpe, 2018; Xia, Manassi, Nakayama, Zipser & Whitney, 2020), and observers readily recognize peripheral faces (McKone, 2004; Farzin et al., 2009) and make accurate saccades to them (Farzin, Rivera & Whitney, 2010; Crouzet et al., 2010; Martin et al., 2018; Wolfe & Whitney, 2015). We foveate faces at least in part because face recognition improves with access to fine detail (Goffaux, Hault, Michel, Vuong & Rossion, 2005; Kwon & Legge, 2011) and for other reasons like lip reading (Lansing & McConkie, 2003; Schwartz, Berthommier & Savariaux, 2004). As a consequence, humans learn to recognize faces by foveating them, and we generally have more experience foveating faces. This could potentially lead to stronger or more numerous holistic templates in the central visual field compared to the peripheral visual field.

The foveal tuning for holistic processing found here is consistent with the extensive brain imaging literature that shows a central visual-field bias for faces which does not exist for other objects (Amedi et al., 2001; Baldassano et al., 2012; Grill-Spector & Weiner, 2014; Hasson et al., 2002; Kamps et al., 2020; Kanwisher, 2001; Levy et al., 2001; Malach et al., 2002; Sayres & Grill-Spector, 2008; Silson et al., 2016). That is, face representations are more affected by eccentricity than other objects (Kreichman et al., 2020; Schwarzlose et al., 2008). Our results also support this retinotopic organization of face processing (Hasson et al., 2002; Levy et al., 2001), but they also extend and clarify it, revealing an explanation for the reported foveal bias in neuroimaging results. In particular, our results help explain

how commonly used neuroimaging methods reveal and reinforce the apparent foveal bias in face representations. For example, neuroimaging studies using univariate and multivariate approaches commonly contrast or classify the differences between object and face responses (e.g., using multivariate pattern analysis to classify faces versus objects). Comparing (implicitly or explicitly) responses to upright faces relative to either objects or inverted faces could create a bias that more heavily weighs holistic processes. Our psychophysical results show that these holistic processes are relatively tuned toward the central visual field. This naturally leads neuroimaging studies to estimate (and potentially overestimate) the foveal tuning of face processing. Our results suggest that the central-field bias does, in fact, reflect a spatial inhomogeneity in face processing—a modest but consistent strength in the holistic representations of faces.

Our results help reconcile the debate on the eccentricity-dependence of face perception, and they make a prediction: faces that are processed by parts (fractured faces, or two-tone faces that maintain featural information, for example) will not result in the kind of foveal bias that the neuroimaging literature has reported. Additionally, those fractured faces will not have such a strong foveal bias when measured psychophysically either. In contrast, faces that require holistic processing, such as hard Mooney faces, will show stronger tuning to the fovea.

Within-face crowding has been shown to impair recognition of gray-scale face features in the periphery (Kalpadakis-Smith, Goffaux & Greenwood, 2018; Martelli, Majaj & Pelli, 2005; Liu, Montaser-Kouhsari & Xu, 2014). Our results show that Mooney faces (especially relatively low-holistic Mooney faces) are well-recognized across the visual field. From this and previous work on peripheral recognition of Mooney faces (Farzin et al., 2009; Farzin et al., 2010; Goold & Meng, 2016), it may be that Mooney faces do not suffer within-face crowding to the same extent that grayscale faces do. We propose two reasons to explain the weaker within-face crowding in our experiment. First, the task tested here required discerning the gender of two-tone faces, which is a demanding task that involves recognizing global and configural cues and relies on holistic processing, especially for those high holistic Mooney faces that isolate holistic processing (Yokoyama et al., 2014; Zhao & Hayward, 2010). Within-face crowding has only been shown for single feature discrimination tasks, like recognizing the shape of the mouth or eyes (Martelli et al., 2005; Liu et al., 2014). For true holistic tasks (identity and gender recognition of Mooney faces, for example), crowding also happens between whole faces (Farzin et al., 2009; Louie et al., 2007). Second, the weaker within-face crowding in Mooney faces may support two parallel ways to recognize the gender of a face. Even if there were crowding between the features of the face (thus impairing recognition of a particular eye or nose), there could still be sufficient holistic information about the structure of the face or its identity that could give away information about the gender of the face (e.g., Fischer & Whitney, 2011).

Holistic and part-based perception have been defined as two independent systems that can operate simultaneously and in parallel (McKone, 2004; Moscovitch et al., 1997). The redundancy in face processing pathways has several advantages. These include 1) a degree of resilience to brain damage (Boutsen & Humphreys, 2002; Moscovitch et al., 1997; Rivest,

Moscovitch & Black, 2009); 2) a plausible way to release crowding (Manassi & Whitney, 2018); and 3) increased resources devoted to face recognition in the central visual field while still maintaining robust face detection mechanisms throughout the periphery (Boucart et al., 2016). The redundancy in face recognition processes also results in complex interactions in face recognition across the visual field. A face that isolates holistic processing, such as our high holistic Mooney faces, can only be processed by the holistic system, and its recognition is therefore somewhat easier in the central visual field. A low-holistic face might benefit from both systems (holistic and part-based processes) and thus can be recognized more easily across the visual field, which would explain the overall better gender recognition of low holistic faces in our data.

A limitation of this study is that it does not address task-dependent effects on the tuning of faces across the visual field. It is possible that the dissociation of results found in the neuroimaging and behavioral literature arises in part from a difference in the tasks employed. In our study, we used a gender discrimination task because it allowed us to test a large sample of faces while avoiding “scrambled” face lures that can introduce bias. Previous literature has shown that gender and identity recognition are intimately linked (Baudouin & Tiberghien, 2002; Ganel & Goshen-Gottstein, 2002; Sergent & Signoret, 1992; Zhao & Hayward, 2013). This may be especially true for Mooney faces: Sergent and Signoret (1992) found that prosopagnosic patients show impaired Mooney face identity and gender recognition but not grayscale face gender recognition. That impaired identity recognition and impaired gender recognition are coupled specifically in Mooney faces suggests that a Mooney gender task such as ours likely involves similar mechanisms. Based on the past literature, we have no reason to believe the choice of our task would be an influencing factor in our study. Nevertheless, future research should investigate the importance of the task on face processing across the visual field more directly.

It is worth noting that holistic processes may not be limited to faces (Bukach, Phillips & Gauthier, 2010; Busey & Vanderkolk, 2005; Wong, Palmeri & Gauthier, 2009). This is hotly debated, of course, but to the extent that holistic representations are useful in non-face object and pattern recognition, the results here might apply. This could have consequences in applied settings. For example, in medical image perception, clinicians and pathologists routinely make consequential perceptual decisions about lesions in radiographs and photographs (Waite et al., 2019). These life-and-death decisions are based on pattern and object recognition abilities, but what mechanisms underlie these has remained elusive (Kok, Van Geel, van Merriënboer, & Robben, 2017). Holistic processes might be involved; for example, recent work has suggested that there are some inversion effects in lesion recognition (Chin, Evans, Wolfe, Bowen & Tanaka, 2018; Sheridan & Reingold, 2017), and this inversion effect can be stronger in expert radiologists. This echoes face recognition inversion effects and is a striking finding. It indicates that holistic representations or pathways can be perceptually learned, and that lesion recognition depends on expertise. However, an inversion effect itself does not rule out innate mechanisms. There could be self-selection in career trajectory or selection-bias during training, as well. This might seem fanciful, but face recognition (a kind of pattern recognition) is largely genetically determined (Wilmer et al., 2010). Whether the

same is true for general pattern recognition is unclear. Likewise, whether holistic-based medical image perception is something that can be learned remains an outstanding question. It is therefore important to understand how holistic processing develops, what kind of perceptual learning is required, and what degree of expertise is involved. The approach employed here and elsewhere—using two-tone Mooney images to test for holistic processes—could be useful in medical image perception domains as well.

In sum, our results suggest that there is a foveal tuning in holistic processing of Mooney faces. This finding reconciles the debate in the literature between behavioral studies that suggest relatively unimpaired face perception in the periphery and brain imaging studies that suggest a central-field bias for face representations. The foveal tuning found for faces is specifically stronger for holistic processing. The results support two face recognition pathways that can work in parallel, which, in combination, support the two complementary goals of search and scrutiny: robust face detection throughout the visual field with concurrent highly detailed and precise representations of upright faces in the central visual field.

3.5 Acknowledgements

The authors are grateful for helpful discussions with Ervin Hafer, Ken Nakayama, and William Prinzmetal. The authors also thank Aneesa Conine-Nakano for her assistance in data collection. This work was supported by the National Institutes of Health [RO1CA236793] (DW); and the United States-Spanish Fulbright Commission (TCB).

Chapter 4

Individual Differences in Classification Images of Mooney Faces.

4.1 Abstract

Human face recognition is robust even under conditions of extreme lighting and in situations where there is high noise and uncertainty. Mooney faces are a canonical example of this: Mooney faces are two-tone shadow-defined images that are readily and holistically recognized despite lacking easily segmented face features. Face perception in such impoverished situations—and Mooney face perception in particular—is often thought to be supported by comparing encountered faces to stored templates. In the present study, we used a classification image approach to measure the templates that observers use to recognize Mooney faces. Visualizing these templates reveals the regions and structures of the image that best predict individual observer recognition and they reflect the underlying internal representation of faces. Using this approach, we tested whether there are face templates that are stable from session to session, whether they support the holistic processing of Mooney faces, whether the templates are observer-specific, and whether they allow for pattern completion of holistic representations even in the absence of an underlying signal. We found that classification images of Mooney faces were indeed consistent within each observer, but they were different between observers, indicating that humans have stable and unique face templates that could drive idiosyncratic individual differences in holistic face recognition. Moreover, we found templates reflected information within the blank regions of the original Mooney faces, suggesting that observers filled-in missing information by matching the stimuli to idiosyncratic internal face templates.

4.2 Introduction

Faces play a central role in our everyday life, and our visual system is in turn extremely sensitive to them. Faces convey critical social and emotional information; they guide in-

teractions and our everyday behavior (Costela & Woods, 2019; Coutrot & Guyader, 2014; Jack & Schyns, 2015). Reinforced by a lifelong exposure to faces, humans have developed a remarkable preference for faces over other objects from infancy (Farzin et al., 2009; Frank et al., 2009; Pascalis & Kelly, 2009; Umiltà et al., 1996). We quickly find faces in a scene and direct our attention and gaze towards them (Bindemann et al., 2005; Boucart et al., 2016; Cerf et al., 2009; Cerf et al., 2008; Costela & Woods, 2019; Coutrot & Guyader, 2014; Crouzet et al., 2010; Foulsham et al., 2010; Jack & Schyns, 2015; Marat et al., 2013; Martin et al., 2001; Ro et al., 2001; Theeuwes & Van der Stigchel, 2006; Wolfe & Whitney, 2015; Xia et al., 2020). In fact, humans can recognize faces under a wide range of conditions, including under challenging environments (Burton et al., 1999; Hole et al., 2002). Humans have also developed expert processing mechanisms specific to faces such as holistic processing, the perception of faces as a whole (Farah et al.; Maurer et al., 2002; Sergent, 1984). Faces therefore enjoys a unique and privileged position in human visual processing.

Previous work suggests that face recognition may involve comparing encountered faces to internal face representations – templates (Cavanagh, 1991; Damasio, Damasio & Van Hoesen, 1982; Sekuler & Abrams, 1968; Sergent, 1984; Simpson & Crandall, 1972; Smith & Nielsen, 1970; Valentine, 1991). Faces that best match those templates are processed more efficiently than those that do not (Rossion, 2008; Rossion and Boremanse, 2008). This matching process allows the recovery of 2 and 3D information from the observed faces, even in impoverished or extreme lighting conditions (Cavanagh, 1991). Valentine (1991) defined the norm-based coding model, which proposes that faces are encoded as distance vectors to a prototype shaped by individual prior experience. Others have aimed to define the structures of these templates and proposed similar models of face recognition that occur by matching observed faces to template-like structures (e.g., Bar codes: Dakin & Watt, 2009; Integral Images: Viola & Jones, 2004; Face Recognition Units: Bruce and Young, 1986, 2011). In these models, faces can be encoded in terms of low-level features that can be parsed or easily segmented. Although these models can accommodate grayscale faces well, they may not be as effective for faces that require holistic processing. For example, they would not be as successful for two-tone shadow defined faces (Mooney faces; Fig. 1) because the local features, edges, contours, and spatial frequency content of Mooney faces are not discriminative of the image as a face (Andrews & Schluppeck, 2004; Bona et al., 2016; Cavanagh, 1991; Farzin et al., 2009; Latinus & Taylor, 2005; McKone, 2004; Moore & Cavanagh, 1998). In principle, template-based models are an appropriate way to analyze Mooney faces (Cavanagh, 1991), but the templates need to carry information about the holistic aspects of the face.

Previous work has shown that there are individual differences in holistic processing of both grayscale and two-tone Mooney faces: some individuals process faces more holistically than others (DeGutis, Mercado, Wilmer, & Rosenblatt, 2013; DeGutis, Wilmer, Mercado, Cohan, 2013; Gauthier, 2020; Russell et al., 2009; Wang et al., 2012). Interestingly, there are also stimulus-specific individual differences in holistic processing (Canas-Bajo & Whitney, 2020): specific faces that are processed holistically by one observer are not necessarily processed holistically by other observers. The origin of these differences remains unclear. Here, our

goal was to test whether such idiosyncrasy in holistic processing is caused by underlying individual differences in observer-specific face templates.

Capturing or visualizing templates can be challenging because they are implicit representations. One of the most widely used and established techniques for measuring templates is the reverse-correlation image classification method (see Murray, 2011 for a review). This technique is a data-driven method that has been widely used in the literature to capture internal representations—the image regions and structures that are most important for detection, discrimination, or classification (Ahumada, 1996, 2002; Ahumada & Lovell, 1971; Beard & Ahumada, 1998; Mareschal, Dakin & Bex, 2006; Neri & Heeger, 2002; Ringach & Shapley, 2004; Solomon, 2002; Victor, 2005; Watson, 1998). More recently, the method has also been used to visualize face templates (Brinkman, Todorov & Dotsch, 2017; Dotsch & Todorov, 2012; Dotsch, Wigboldus, Langner & van Knippenberg, 2008; Éthier-Majcher, Joubert & Gosselin, 2013; Gosselin and Schyns, 2003; Karremans, Dotsch, & Corneille, 2011). In a standard reverse-correlation experiment, the stimuli consist of random noise overlaid over the same base face. The observers' task is to classify the stimuli based on some construct (eg., trustworthiness: Éthier-Majcher et al., 2013). Classification images are computed by averaging the noise patterns of the stimuli that participants classified as being representative of the construct of interest. Classification images can therefore serve as a visualization of internal representations of faces (Brinkman, Todorov & Dotsch, 2017; Dotsch & Todorov, 2012; Gosselin & Schyns, 2003; Sekuler, Gaspar, Gold & Bennett, 2004). The reverse-correlation technique is ideal to tap into implicit representations of faces because it is a purely data-driven method: noise is generated randomly in each trial for each participant, and participants use their own internal representation of the construct without external biases (Mangini & Biederman, 2004; Todorov, Dotsch, Wigboldus, & Said, 2011). In the field of face recognition, this method has been used to understand the basis of face recognition at the group level (Mangini & Biederman, 2004; Sekuler et al., 2004), and how it is influenced by factors like emotion (Brooks & Freeman, 2018; Karremans et al., 2011; Jack, Caldara & Schyns, 2012), perceptual disorders (Brinkman et al., 2019), prejudice (Dotsch, Wigboldus, & Van Knippenberg, 2013), social cognition (Brooks, Stoller & Freeman, 2018) and culture (Dotsch et al., 2008; Jack et al., 2012; Ratner, Dotsch, Wigboldus, van Knippenberg, & Amodio, 2014). For instance, Dotsch and Todorov (2012) used classification images to identify racial biases in the perception of trustworthiness in faces.

Here, we use the reverse correlation technique to investigate whether there are individual differences in classification images of Mooney faces. Mooney faces are highly impoverished two-tone black and white blobs that are readily perceived as faces despite lacking low-level face features that can be parsed in a bottom-up fashion (Fig. 1; Brodski et al., 2015; Gould and Meng, 2016; Ke et al., 2017; Mooney, 1957; Schwiedrzik et al., 2018). That is, the image must be recognized as a face before any feature (e.g., an eye, a mouth, etc) can be identified or localized (Cavanagh, 1991; Farzin et al., 2009). Consequently, holistic processing plays a critical role in the recognition of Mooney faces. Holistic processing breaks down with inverted faces, which makes inverted Mooney faces very difficult to recognize (Andrews & Schluppeck, 2004; Bona et al., 2016; Cavanagh, 1991; Farzin et al., 2009; Latinus & Taylor,



Figure 4.1: Mooney faces used in this study in upright (top) and inverted (bottom) orientation. Mooney faces are two-tone images that can be quickly recognized as faces despite lacking low-level face features. Mooney faces are processed as a whole; note that while upright Mooney faces are easily perceived as faces, inverted Mooney faces are difficult or even impossible to recognize, as such. For this reason, Mooney faces are ideal stimuli to test holistic processing.

2005; McKone, 2004; Moore & Cavanagh, 1998; Sargent, 1984). For these reasons, Mooney faces are ideal stimuli to isolate holistic processing (Fig. 1).

The following experiments tested several hypotheses. First, because our goal was to measure the templates that support holistic perception of Mooney faces, we expected that classification images would be consistent from day to day. Second, if previously found individual differences in holistic processing reflect underlying observer-specific templates (Canas-Bajo & Whitney, 2020), we should find idiosyncratic classification images only for upright but not inverted Mooney faces. Lastly, if Mooney faces are recognized by matching the stimuli to stored templates, classification images of Mooney faces should reflect filled-in information that is missing in the original, impoverished stimuli. Altogether, our study aimed to investigate whether there are observer-specific face templates (Experiment 1) and whether they allow for pattern completion of the missing information in the original Mooney faces (Experiment 2).

4.3 Experiment 1

4.3.1 Methods

4.3.1.1 Participants

7 participants (3 male, 4 female) took part in this experiment. One participant was excluded because they did not complete all trials of the experiment. All subjects were recruited through Mechanical Turk and provided written consent form before participation. All experimental procedures were approved by the UC Berkeley Institutional Review Board.

4.3.1.2 Material & Design

The base faces in the reverse correlation task consisted of four unfamiliar Mooney faces created by Schwiedrzik et al. (2018). These Mooney faces were presented both in upright and inverted orientations, generating a total of eight base faces for the reverse correlation experiment (Fig. 1). In order to assess the consistency of the subjects' classification images, participants did the experiment twice in two separate days. The time interval between the two sessions was on average 3.4 days (std: 1.5 days). In each experimental session, participants completed 8 blocks of 250 trials for each base image.

The presentation of the stimuli and the data collection was controlled using Qualtrics (<https://www.qualtrics.com>). Note that participants used different monitors, so distance to monitor could not be controlled.

4.3.1.3 Procedure

Stimuli were generated uniquely for each participant. In each trial, we generated a random pattern of sinusoidal noise and its corresponding negative version (Figure 2A; for more details on how the random noise was generated, see Dostch & Todorov, 2012). Each trial consisted of two side-by-side stimuli: the base face with the original noise superimposed and the same base face with the polarity reversed noise superimposed (Fig. 2C). Participants' task was to select which of the two images presented was more face-like by pressing the left or right buttons. The stimuli were presented until observers gave a response. Participants could move their eyes freely during the trial and could take a break every 250 trials. The patterns of noise were randomized for each trial separately for each observer. That is, no two subjects saw the same stimuli, and the same observer did not see the same stimuli more than once. All participants completed a demographics questionnaire followed by a question regarding their previous experience with Mooney faces.

4.3.1.4 Data Analysis

Our goal here was to investigate whether face templates support recognition of Mooney faces, and whether the templates, if present, are observer-specific. We used classification images to

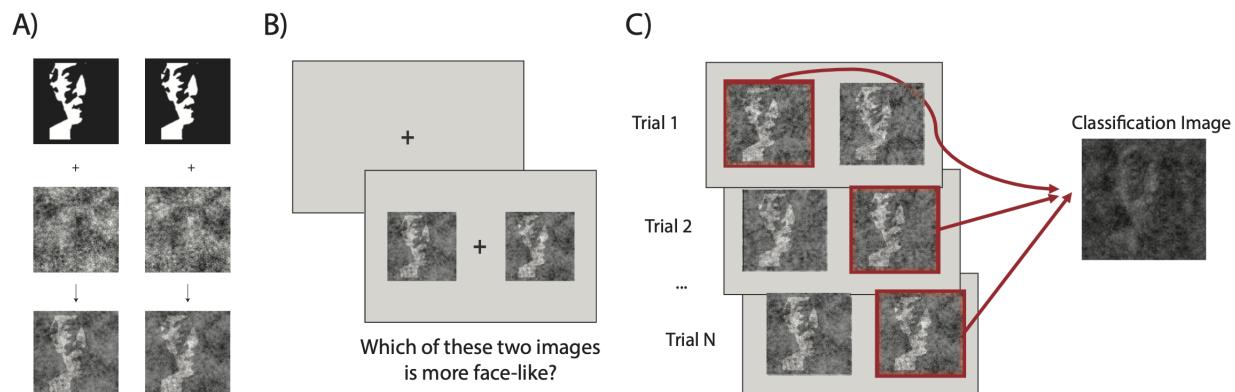


Figure 4.2: To generate the stimuli for each trial, the same base image was overlaid with two identical, but polarity-reversed random patterns of noise. B) Example trial. After a blank screen, two stimuli were presented, and participants were asked to select which of two images was more face-like. C) Classification images were generated by averaging the selected noise patterns across trials.

visualize observers' face templates. To generate the classification images, we averaged all of the observer's selected noise patterns across trials, and we superimposed the resulting noise on the base face (Fig. 2C). The classification image reveals the regions or structures that observers rely on to classify an image as face-like. We generated one classification image per base face, per subject. The outcome was eight unique classification images for each observer.

To evaluate if observers' classification images were stable over time, we tested the day-to-day consistency of each observers' classification images for each of the 8 base Mooney faces. The within-subject consistency was defined as the test-retest pixel-wise correlation between an observer's classification images in the first and second session. We then averaged all within-observer Fisher z transformed correlations for upright and inverted base faces, separately. Second, we quantified the between-observer agreement in classification images of Mooney faces. To this end, we calculated the between-observer correlations across different observers' classification images for each base face. Then, we averaged all between-observer Fisher z transformed correlations for upright and inverted base faces, separately.

To test the significance of the within and between-subject correlations, we calculated a null distribution of permuted correlations. In each iteration of the permutation, for each base face, we shuffled the responses that were given by each participant. We then calculated a classification image from these shuffled responses. This was effectively the classification images an observer would have if the participant responded randomly. Using this data, we generated a set of 1000 permuted classification images, per base face, per participant. Note that our classification images have signal included, which could artificially inflate the correlations in both within and between subject analyses. To control for this, we calculated null distributions of permuted correlations, which maintain all of the same signal information,

but represent shuffled responses, so any inflation of the correlations will occur in the null distribution as well. In other words, we never compare correlations to zero, always to permuted shuffled null distributions. To generate the distribution of permuted null within-subject correlations, for each observer and each base face, we correlated the empirical classification image of the first session with each of the permuted null classification images of the second session. In the same fashion, to generate the distribution of permuted null between-subject correlations, for each pair observers and each base face, we correlated an observer’s one-session empirical classification image with each of the permuted null classification images of another observer’s session.

Lastly, we quantified the individual differences of classification images of Mooney faces by comparing within and between-observer agreement. The significance of this comparison was tested using a non-parametric permutation method. Across all analyses, to calculate the average within and between-subject correlations across base faces and/or across observers, correlations were first transformed from Pearson R correlations into Fisher z correlations before averaging them.

4.3.2 Results

In this study, we measured classification images of Mooney faces to investigate whether there are face templates that support holistic face recognition and whether these templates were observer-specific. In our experiment, participants classified which stimuli were more face-like.

The outcome per observer was a classification image for each base face tested (see Fig. 3 for some examples). Casual inspection hints that the classification images of upright faces seem to be visible and face-like (Fig. 3A). In contrast, classification images of inverted faces appeared to be less recognizable (Fig. 3B). Our hypothesis was that an observer’s classification images represented the observer’s underlying face representations. To test whether classification images were even above chance, we collapsed the upright and inverted faces (vertical pairs of faces in Fig 3A-B) together. The goal of this analysis was to check and confirm that we could even measure classification images that were significantly different than random responses. If classification images of face-likeness can be measured, then there should be significant within and between subject consistency. We found that there was both a significant within and between-observer correlation for collapsed classification images of Mooney faces (within: Fisher $Z = 0.22$, $p < 0.01$; between: Fisher $Z = 0.21$, $p < 0.01$; Fig. 4A). This result shows that we effectively measured classification images of face-likeness and that they were not due to random responses.

The Mooney faces that we used in this experiment require holistic processing to be recognized (Canas-Bajo & Whitney, 2020; 2022; Farzin et al., 2009; Latinus & Taylor, 2005; McKone, 2004; Moore & Cavanagh, 1998). If recognition of Mooney faces is supported by holistic face templates, then classification images of Mooney faces should reflect an advantage for upright faces. Inverted Mooney faces are hard or even impossible to recognize (Kanwisher et al., 1998; Latinus & Taylor, 2005; McKone, 2004), so we expected weaker classification images of inverted Mooney faces. Figure 3 seems to support this division, showing what appear

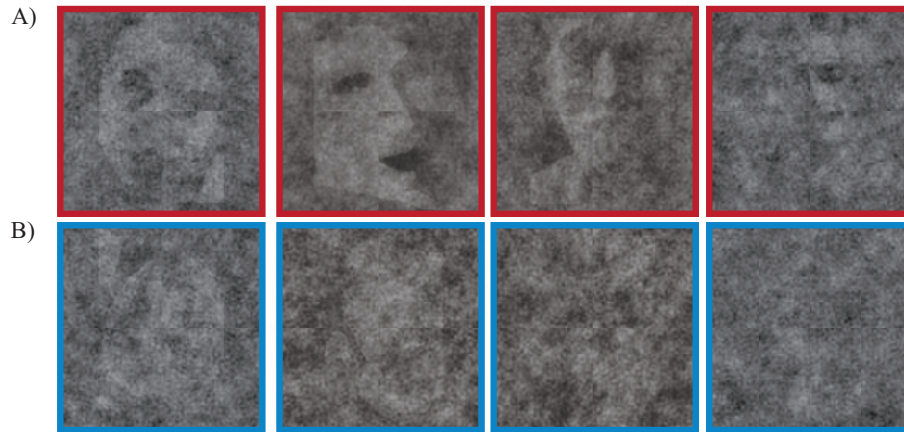


Figure 4.3: Sample resulting classification images of Mooney faces in Experiment 1. A) Top row: upright Mooney faces (in red). The Mooney face Classification images are recognizable (albeit noisy and ghostly) B) Bottom row: inverted Mooney faces (in blue). Inverted Mooney face classification images did not appear to be very obvious. See main text for details.

to be weaker classification images for inverted faces. To address this question quantitatively, we computed the within- and between-observer agreement for classification images of upright and inverted faces, separately. We found that there was a significant within-observer agreement in classification images for upright faces (Fisher $Z = 0.26$, $p < 0.001$, per permutation test; Fig. 4B) but not inverted faces (Fisher $Z = 0.18$, $p > 0.05$, per permutation test; Fig. 4C). That is, only upright Mooney base faces led to stable classification images from session to session. Inverted Mooney faces did not. Similarly, we found a significant between-observer agreement only for upright Mooney faces (Fisher $Z = 0.23$, $p < 0.01$; Fig. 4B), but not inverted Mooney faces (Fisher $Z = 0.19$, $p > 0.05$; Fig. 4C).

One of our primary goals was to investigate whether classification images of Mooney faces were observer-specific. To this end, we compared the within- and between- observer agreement in the classification images of upright and inverted faces, separately. We found that within-observer agreement of classification images of upright Mooney faces was significantly higher than the between-observer agreement ($p < 0.05$, per permutation test; Fig. 4B). This difference was not significant in classification images of inverted Mooney faces ($p > 0.05$, per permutation test; Fig. 4C). This result suggests that there are stable individual differences in the classification images of upright Mooney faces, which require holistic processing.

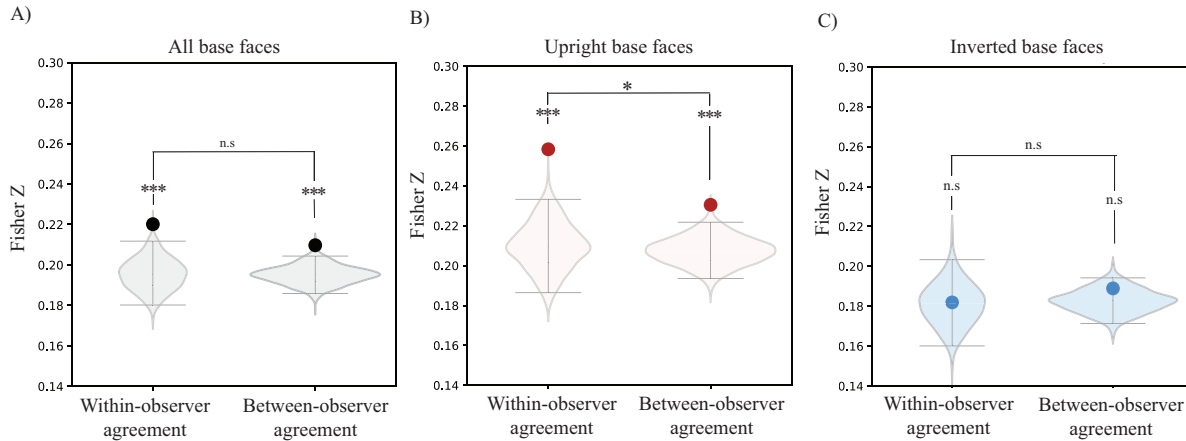


Figure 4.4: Results of Experiment 1. A) Within and between-observer agreement for classification images collapsed across upright and inverted faces. B) Within and between-observer agreement for classification images of upright base Mooney faces. C) Within and between-observer agreement for classification images of inverted base Mooney faces. In A-B, solid dots represent the empirical correlations, and the violinplots represent the respective permuted null distributions. Error bars represent the upper 97.5% and lower 2.5% boundaries of the permuted null correlation distribution. In both panels, the stars represent the significance of the permutation test. Three stars represent $p < 0.001$, two stars represent $p < 0.01$ and one star represents $p < 0.05$. Non-significant comparisons are indicated by n.s.

4.3.3 Discussion

The results of Experiment 1 indicate that classification images of Mooney faces are stable from session to session. We also found that there are individual differences in classification images of Mooney faces, specific to upright faces. This result is consistent with the idea that holistic face recognition is supported by matching Mooney faces to face templates (Cavanagh, 1991) and that these templates are idiosyncratic—unique to each individual observer. One concern with Experiment 1 is that data were collected online using the Qualtrics platform, and participants completed the experiment on their own computer monitors. Thus, we could not gamma correct or control the linearity of each individual participant’s monitor. Consequently, the individual differences found may have been influenced by differences across monitors. It is important to note that our inverted condition served as a partial control for this possibility, and the dissociation in results found for upright and inverted faces suggests that the individual differences found in upright Mooney faces are probably not due to differences in gamma correction across the monitors. Nevertheless, individual differences in monitors are an important concern, as are individual differences in room set up, ambient lighting, monitor distance, and other environmental factors. We therefore conducted a sec-

ond experiment, in the laboratory, with a single controlled and calibrated monitor. The goal of Experiment 2 was to replicate Experiment 1 results and to further investigate whether the individual differences in classification images reflect an underlying idiosyncrasy in the internal representations of faces.

4.4 Experiment 2

Experiment 2 had two goals. First, to control for a possible contribution of individual differences in monitors, hardware, software settings, environmental factors, or any other differences due to the online nature of the first experiment. We therefore conducted the experiment in-person with a controlled environment, constant hardware, and gamma correction of the monitor used for all subjects. The second, more theoretically important goal of Experiment 2 was to investigate whether there is pattern completion within classification images of Mooney faces. If Mooney faces are recognized through a matching process to stored face templates, then classification images should show that observers fill in missing face feature information within the original Mooney face. Additionally, if holistic face recognition is supported by idiosyncratic templates as suggested by Experiment 1, then we should expect that the way information is filled in is unique to each observer, at least for upright faces.

4.4.1 Method

4.4.1.1 Participants

Nineteen participants (8 male, 11 female) took part in this experiment. All subjects were undergraduate students at the University of California, Berkeley and provided written consent form before participation. All experimental procedures were approved by the UC Berkeley Institutional Review Board.

4.4.1.2 Material

Stimuli were presented on a gamma corrected CRT monitor at 100 Hz refresh rate, with 1024x768 pixels resolution and a horizontal screen size of 40.5cm. The monitor was placed 60cm from a chin rest that stabilized the participant's head. At this distance, all the face stimuli shown during the experiment subtended 6 degrees visual angle. The presentation of the stimuli was controlled using MATLAB R2016b with Psychophysics Toolbox 3 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007).

4.4.1.3 Design, procedure and data analysis

Design, procedure, and data analysis was the same as in Experiment 1. Average time between participants' experimental sessions was 1.7 days (STD = 0.8)

4.4.2 Results

The first goal of Experiment 2 was to replicate Experiment 1’s results with a single gamma-controlled monitor and consistent environment across observers. Consistent with Experiment 1, we found significant within and between observer agreement in classification images collapsed across upright and inverted conditions, which further confirms that that classification images of Mooney faces are stable from session to session and not due to random noise (within correlation: Fisher $Z = 0.23$, $p < 0.001$; between-subject correlation: Fisher $Z = 0.22$, $p < 0.001$; Fig. 5A)

We next compared the consistency and uniqueness of classification images of upright and inverted faces, separately. As in Experiment 1, classification images of upright faces displayed a significant within and between observer agreement (within-observer Fisher $Z = 0.24$, $p < 0.001$; between-observer Fisher $Z = 0.22$, $p < 0.001$; Fig. 5B). In this experiment, we found that the within and between-observer agreement remained significant for inverted faces (within-observer Fisher $Z = 0.22$, $p < 0.001$; between-observer Fisher $Z = 0.22$, $p < 0.001$, per permutation test; Fig. 5C), but the day-to-day consistency was nevertheless weaker than in classification images of upright faces.

Importantly, we found a relatively higher within-observer agreement than between-observer agreement that was specific to upright faces ($p < 0.05$, per permutation test; Fig. 5B). Replicating Experiment 1, this finding indicates that there are individual differences in classification images that are specific to upright Mooney faces, which require holistic processing. The individual differences found in Experiment 1 are evidently not due to the lack of a consistent environment across observers. In contrast, classification images of inverted faces were not found to show observer-specific idiosyncrasies ($p > 0.05$, per permutation test; Fig. 5C).

So far, we showed that there are individual differences in classification images of Mooney faces that are consistent session to session. However, the question still remained as to what these day-to-day consistent classification images represent. Our hypothesis was that they reflected participants’ internal representation of faces. Mooney faces do not contain easily parsed or bottom-up segmentable information about the parts of the faces (one has to know it is a face before localizing or identifying a particular eye or nose), and some or all face features can be missing from them (e.g., a half-illuminated faces). If classification images reflect participants’ templates for faces, then they might reveal filled-in information that is missing at the image level. For example, a base Mooney face that is only partially illuminated on the left side, and mostly in complete darkness on the right side (Fig. 6A), might nevertheless evoke a classification image that includes details in the invisible portion of the image. To test this hypothesis, we investigated whether there is pattern completion in classification images of Mooney faces. We operationalized pattern completion as the information present in the classification image only within the black regions of the Mooney face (regions highlighted in red in Fig. 6B). To increase the power of the pattern completion analysis and considering the similarity in results of Experiment 1 and 2, we collapsed participant data from both experiments for the following tests.

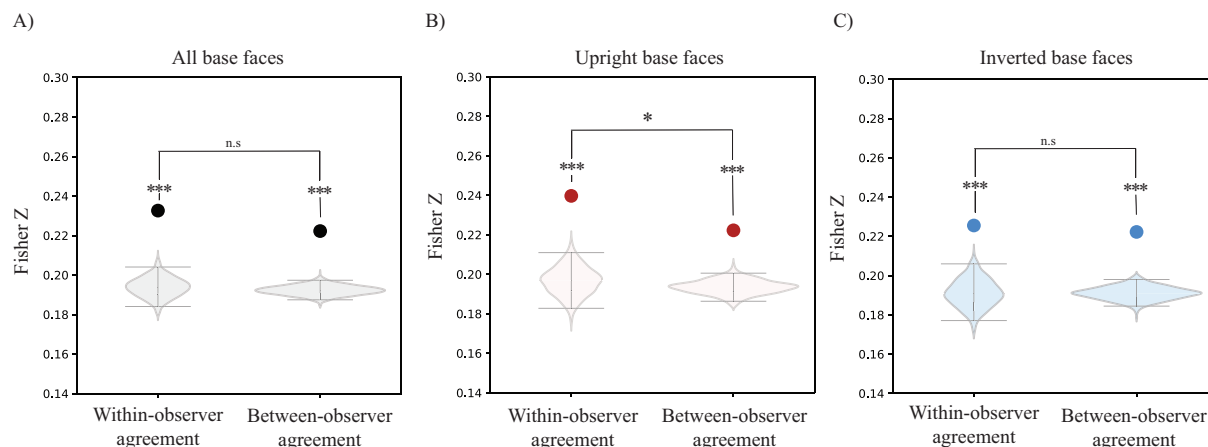


Figure 4.5: Results of Experiment 2. A) Within and between-observer agreement for classification images of all upright and inverted base faces collapsed. B) Within and between-observer agreement for classification images of upright base faces. C) Within and between-observer agreement for classification images of inverted base faces. In A-B, black dots represent the empirical correlation found and the violin plots represent the respective permuted null correlations. Error bars represent the upper 97.5% and lower 2.5% boundaries of the permuted null correlation distribution. In both panels, the stars represent the significance of the permutation test. Three stars represent $p < 0.001$, two stars represent $p < 0.01$ and one star represents $p < 0.05$. Non-significant comparisons are represented with n.s.

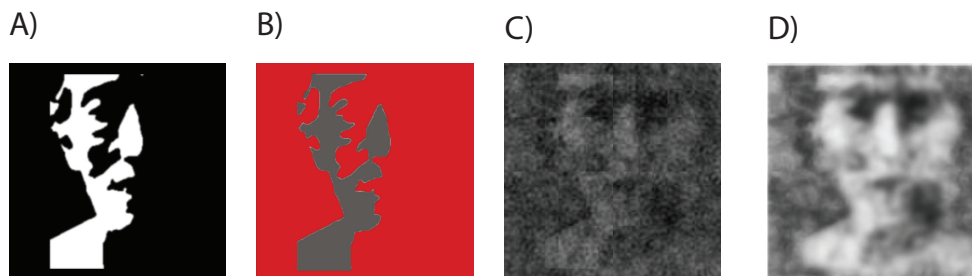


Figure 4.6: A) Original Mooney face used as base face. B) In red, original black areas in classification images. Correlations calculated within these signal-free and otherwise featureless regions to analyze the extent of pattern completion. C) Classification image reflecting pattern completion. D) Classification image with enhanced contrast to facilitate visualization of the filled-in areas.

To measure if classification images showed pattern completion of missing face features within the black regions of the original Mooney face, we calculated the within-observer agreement within these black regions (Fig. 7A). These black regions do not contain any signal, so any consistency in the selected noise patterns within these areas would suggest that the observer has perceptually completed parts of the face that were not originally there. We found significant within and between observer agreement within the black regions of all collapsed classification images (within-observer correlation: Fisher $Z = 0.03$, $p < 0.001$; between-observer correlation: Fisher $Z = 0.02$, $p < 0.001$; Fig. 7A). This result indicates there is pattern completion in classification images of Mooney faces and corroborates the hypothesis that recognition of Mooney faces is supported by a matching process to stored templates (Cavanagh, 1991).

When we explored pattern completion within classification images of upright and inverted Mooney faces separately, we found a significant within-subject correlation within the black regions of upright Mooney faces (Fisher $Z = 0.04$, $p < 0.001$, per permutation analysis; Fig. 7B). This indicates that there was unique filling-in of missing information within single individual observers. Moreover, this pattern completion was consistent from session to session. This confirms that the pattern completion result found in upright Mooney faces was not random; that is, there was meaningful information present in those regions that was not originally there. As a visualization, note a representative observers' classification image in which pattern completion is evident (Fig. 6): even though there is no single face feature information on the right side of the original upright Mooney face (Fig. 6A), the classification image of a sample observer showed qualitative completion of the missing features (the right eye, the chin line, the continuation of the mouth, etc.; Fig. 6C). Pattern completion becomes more evident visually when we enhance the contrast of the image (Fig. 6D). We found no consistent pattern completion for inverted Mooney faces (Fisher $Z = 0.02$, $p > 0.05$; Fig. 7C).

Moreover, if observers' templates that support face recognition are idiosyncratic, then classification images should show unique pattern completion for each participant. We found a significant, but relatively low between-observer agreement in the way black regions in upright Mooney faces were filled-in (Fisher $Z = 0.02$, $p < 0.001$; Fig. 7B). This indicates some consistency between observers, but the within-observer correlation was significantly higher than the between-observer correlation ($p < 0.5$, permutation analysis; Fig. 7B). These results suggest that there is pattern completion in high-level object representations of upright faces, and that the perceptually filled-in information was unique to each observer. Inverted faces were different. Although we found significant between-observer agreement in the pattern completion of inverted Mooney faces (Fisher $Z = 0.01$, $p < 0.001$; Fig. 7C), the difference between within and between-observer agreement was not significant. This result suggests that there is weaker and less idiosyncratic pattern completion for inverted faces.

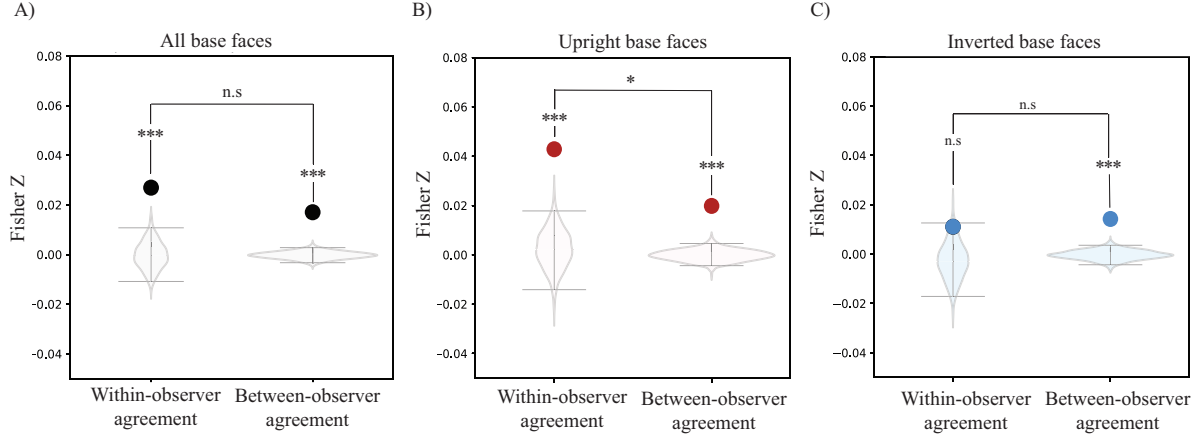


Figure 4.7: Pattern completion results. A) Within and between observer agreement in signal-free black areas of all original upright and inverted Mooney faces collapsed. B) Within and between observer agreement in black areas of the original upright Mooney face. C) Within and between observer agreement in black areas of the original inverted Mooney face. In all panels, black dots represent the empirical correlation found and the violin plots represent the respective permuted null correlations. Error bars represent the upper 97.5% and lower 2.5% boundaries of the permuted null correlation distribution. In both panels, the stars represent the significance of the permutation test. Three stars represent $p < 0.001$, two stars represent $p < 0.01$ and one star represents $p < 0.05$. Non-significant comparisons are represented with n.s.

4.4.3 Discussion

In Experiment 2, we aimed to replicate the results found in Experiment 1 under a more controlled experiment set up. In sum, we found that classification images of Mooney faces were stable from session to session, as in Experiment 1. Similarly, we found observer-specific differences in classification images of Mooney faces that were specific to upright faces. This finding confirms that the results of Experiment 1 were not due to a floor effect.

Furthermore, to understand whether Mooney face recognition is supported by matching them to stored face templates, we explored whether there was stable and unique filling-in of missing facial information in classification images. To address this question, we looked at the within and between-subject agreement only within the black regions of the original Mooney face. We found within-observer agreement within these black regions, indicating that there is pattern completion in high-level object representations. This result suggests that Mooney face perception is supported by matching to stored templates. Pattern completion was stronger in upright than inverted classification images. Interestingly, the filled-in pattern was unique to each observer only in upright faces, which suggests that the idiosyncrasy in face recognition is specific to holistic processing.

4.5 General discussion

Previous work suggests that face recognition is supported by a matching process between observed faces and internal face templates (Cavanagh, 1991; Damasio et al., 1982); Sekuler & Abrams, 1968; Sergent, 1984; Simpson & Crandall, 1972; Smith & Nielsen, 1970; Valentine, 1991). Faces that best match those internal face representations are processed more efficiently than those that do not (Rossion, 2008; Rossion and Boremanse, 2008). However, proposed definitions of face templates often describe a representation of low-level features, which do not generally account for holistic processing of faces and the advantage it affords face recognition (Dakin & Watt, 2009; Viola & Jones, 2004). In the present study, we used classification images of Mooney faces to investigate whether there are face templates that are stable session to session, whether they support holistic processing of Mooney faces, whether they are observer-specific, and whether they allow for pattern completion of holistic representations.

Our results show that there are day-to-day consistent classification images of Mooney faces. The reverse-correlation task used here was purely data-driven: observers were not biased towards any external definition of face-likeness, so they applied their internal face presentations to classify the stimuli as face-like. This suggests that observers' classification images reflected their underlying, implicit face representations. Importantly, we found that classification images of upright Mooney, but not inverted Mooney faces, were observer-specific. This suggests that there are idiosyncrasies in the stored face templates that support holistic processing, which could be driving the individual differences in holistic processing found previously in the literature (Canas-Bajo & Whitney, 2020; de Heering & Rossion, 2008; Goodman et al., 2007; Ferguson, Kulkofsky, Cashon & Casasola, 2009; Kuefner, Cassia, Vescovo & Picozzi, 2010; Susilo, Crookes, McKone, & Turner, 2009). These results support the conclusion that holistic and part-based processing are distinct mechanisms of face recognition (McKone, 2004; Moscovitch et al., 1997), and that holistic processing is idiosyncratic while part-based processing is relatively universal (Canas-Bajo & Whitney, 2020).

Furthermore, we found that observers filled-in missing information in the original base Mooney face. Humans are highly sensitive to incomplete and partial visual information, resulting in object and scene recognition that proceeds seemingly unimpaired by occlusions and noise. Many examples of this have been reported, including contour and surface-based filling-in and pattern completion (Gold, Murray, Bennett, Sekuler, 2000; Sekuler, Gold, Murray & Bennett, 2000). Here, we used Mooney faces as stimuli, which do not contain enough low-level information to be uniquely processed in a part-based manner (Canas-Bajo & Whitney, 2020). The pattern completion results found in the present study cannot be explained by known contour integration, surface filling-in, or feature-based pattern completion processes. Interestingly, we found that the way missing information was filled in was idiosyncratic only in upright Mooney faces. Mooney face recognition may therefore be supported by matching the impoverished images to idiosyncratic stored templates.

Lastly, we also found weaker but stable classification images of inverted Mooney faces with

weaker pattern completion and no idiosyncrasies. The stability of these inverted classification images was less efficient than for upright faces, which confirms that our results were not due to a floor effect in the inverted condition. It also agrees with the idea that the part-based pathway to face recognition may be less efficient (Sekuler et al., 2004). Mooney faces can occasionally maintain some low-level features and be partially supported by a combination of holistic and part-based processing (Canas-Bajo & Whitney, 2020). This suggests that face templates may not only support holistic processing of upright faces but also part-based processing of inverted faces and perhaps object recognition more generally. This is in line with previous work that reported templates for objects (Biederman & Bar, 1999; Rock & DiVita, 1987; Poggio & Edelman, 1990; Logothetis, Pauls, Bülthoff, & Poggio, 1994; Tarr, 1995; Tarr & Bülthoff, 1995).

A long history of computer vision research has aimed to reproduce the template-matching theory of face recognition. Turk and Pentland (1991) defined a computational face recognition system in which faces are encoded as vectors of weights, called eigenfaces, and stored in memory. Eigenfaces are a low-cost representation of faces. When the system encounters a novel face, recognition occurs through a matching procedure to the stored eigenfaces – a type of template. Others have developed Deep Convolutional Neural Networks models of face recognition that highlight the need for internal representations to be general and flexible, so that they can achieve the level of expertise with familiar and unfamiliar faces that we see in humans (Blauch, Behrmann & Plaut, 2021; Orru, Marcialis & Roli, 2020), parallel to the well-known advantage for familiar faces. This and other computer-vision based face recognition research suggests that models that use template-matching can, at least in principle, reproduce the patterns of human face recognition.

Future research can use our approach to investigate the origin and development of face templates at the observer level. The proposed approach of studying holistic perception using reverse correlation classification image method is powerful and could be especially useful in narrowing and identifying the neural mechanisms of holistic face perception. Unfortunately, it is a very time-consuming method, but it can be done within individual observers, as our data suggests. Future work might therefore be able to employ the technique in combination with EEG, fMRI, and MEG. The technique might also be useful for evaluating the biological plausibility and correspondence between face identification algorithms and holistic face perception by assessing classification images in both domains.

Altogether, our results provide evidence that classification images can be used as an unbiased method to tap into observer-specific internal representations of faces. Furthermore, the present study provides a new method for future research to investigate the nature of face templates, such as understanding how many templates are required to support our expert face recognition system (Laurence, Baker, Proietti & Mondloch, 2021); investigating the plasticity of face templates over development (de Heering & Rossion, 2008; Laurence et al., 2021); or making predictions from an observer’s classification image of which faces will be more efficiently processed.

4.6 Acknowledgments

The authors are grateful for helpful discussions with Ervin Hafter, Ken Nakayama, and William Prinzmetal. The authors also thank Maya Zhu, Aidan Gauper and Ethan Shedd for their assistance in data collection. This work was supported in part by United States-Spanish Fulbright Commission and NIH grant RO1CA236793.

Chapter 5

Conclusion

The research presented in this dissertation aimed to understand whether there are individual differences in holistic face perception and whether those differences reflect underlying differences in face templates. Across the dissertation, we used Mooney faces as stimuli because they require holistic processing to be recognized as faces. The findings of these studies show that holistic face recognition is stimuli-specific, observer-specific and tuned to the fovea, and that these differences reflect underlying idiosyncratic face templates.

In Chapter 2, our goal was to investigate whether there are stimuli-specific individual differences in holistic processing of Mooney faces. First, we found that Mooney faces vary in the extent they are processed holistically, which shows that holistic and part-based perception are two independent systems that can operate simultaneously and in parallel (McKone, 2004; Moscovitch, Winocur & Behrmann, 1997). Importantly, stimuli-specific individual differences in recognition of Mooney faces; that is, we found that specific faces that are processed holistically by one observer are not by other observers. Interestingly, we found that there were no idiosyncrasies in recognition of the Mooney faces that could be recognized as a combination of part-based and holistic processing. The results from this chapter indicate that the tuning of holistic processing is observer specific while part-based processing of faces is universally shared across observers.

In Chapter 3, we aimed to investigate whether holistic processing is also visual-field-specific. We observed that holistic faces were foveally tuned, which means that holistic processing was slightly but reliably stronger at the fovea than in the periphery. We found that there was no tuning in Mooney faces that could be recognized as a combination of part-based and holistic processing. Our results suggest that because we most often direct our attention and gaze towards faces in a scene (Costela & Woods, 2019; Jack & Schyns, 2015; Coutrot & Guyader, 2014), faces that appear at the fovea may enjoy a relative advantage in face-specific mechanisms such as holistic processing.

In Chapter 4, we used a classification image approach to investigate whether there are observer-specific differences in face templates that underlie the idiosyncratic differences in recognition found in the previous chapters. Previous work suggests that face recognition is supported by matching faces to stored face templates. However, thus far, proposed templates

could not account for holistic processing of Mooney faces or for the individual differences that we showed here. we used a classification image approach to tap into internal face templates. We found that classification images reflect underlying internal face representations – templates and that they can support holistic processing of Mooney faces. Importantly, we found that classification images enjoyed pattern completion—there was filling-in of missing information in Mooney faces. This further corroborates the idea that matching observed faces to internal face templates is necessary for holistic face perception.

Face recognition is crucial to our daily lives and therefore enjoys a unique and privilege role in cognition, including having a face-specific type of recognition – holistic processing. The goal of this dissertation was to characterize holistic processing and specifically to investigate whether are individual differences in holistic processing of Mooney faces. In sum, the research discussed in this dissertation suggests that (1) holistic face perception is both individual observer and individual stimulus specific, (2) that holistic face perception is tuned to the fovea where faces are most often processed, (3) that observers have face templates that support face recognition through a matching process that fills in missing information. Altogether, our results shed light on the existence and nature of individual differences in face recognition, and they highlight the importance of taking into account individual differences in vision research more generally.

Chapter 6

Bibliography

- Ahumada Jr, A. J. (1996). Perceptual classification images from Vernier acuity masked by noise. *Perception*, 25(1_{suppl}), 2 – 2.
- Ahumada, A. J. (2002). Classification image weights and internal noise level estimation. *Journal of Vision*, 2(1), 8-8.
- Ahumada Jr, A., & Lovell, J. (1971). Stimulus features in signal detection. *The Journal of the Acoustical Society of America*, 49(6B), 1751-1756.
- Amedi, A., Malach, R., Hendler, T., Peled, S., & Zohary, E. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nature neuroscience*, 4(3), 324-330.
- Andrews, T. J., & Schluppeck, D. (2004). Neural responses to Mooney images reveal a modular representation of faces in human visual cortex. *Neuroimage*, 21(1), 91-98.
- Baldassano, C., Iordan, M. C., Beck, D. M., & Fei-Fei, L. (2012). Voxel-level functional connectivity using spatial regularization. *NeuroImage*, 63(3), 1099-1106.
- Baudouin, J. Y., & Tiberghien, G. (2002). Gender is a dimension of face recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(2), 362.
- Beard, B. L., & Ahumada Jr, A. J. (1998, July). Technique to extract relevant image features for visual tasks. In *Human vision and electronic imaging III* (Vol. 3299, pp. 79-85). SPIE.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, 57, 289–300.
- Bi, T., Chen, N., Weng, Q., He, D., & Fang, F. (2010). Learning to discriminate face views. *Journal of Neurophysiology*, 104(6), 3305-3311.
- Biederman, I., & Bar, M. (1999). One-shot viewpoint invariance in matching novel objects. *Vision research*, 39(17), 2885-2899.
- Bindemann, M., Burton, A. M., Hooge, I. T., Jenkins, R., & De Haan, E. H. (2005). Faces retain attention. *Psychonomic Bulletin & Review*, 12(6), 1048-1053.
- Blauch, N. M., Behrmann, M., & Plaut, D. C. (2021). Computational insights into human perceptual expertise for familiar and unfamiliar face recognition. *Cognition*, 208, 104341.

- Bona, S., Cattaneo, Z., & Silvanto, J. (2016). Investigating the causal role of rOFA in holistic detection of mooney faces and objects: an fMRI-guided TMS study. *Brain Stimulation*, 9(4), 594-600.
- Boucart, M., Lenoble, Q., Quetelart, J., Szaffarczyk, S., Desprez, P., & Thorpe, S. J. (2016). Finding faces, animals, and vehicles in far peripheral vision. *Journal of vision*, 16(2), 10-10.
- Boutsen, L., & Humphreys, G. W. (2002). Face context interferes with local part processing in a prosopagnosic patient. *Neuropsychologia*, 40(13), 2305-2313.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, 10(4), 433-436.
- Brinkman, L., Dotsch, R., Zondergeld, J., Koevoets, M. G., Aarts, H., & van Haren, N. E. (2019). Visualizing mental representations in schizophrenia patients: A reverse correlation approach. *Schizophrenia Research: Cognition*, 17, 100138.
- Brinkman, L., Todorov, A., & Dotsch, R. (2017). Visualising mental representations: A primer on noise-based reverse correlation in social psychology. *European Review of Social Psychology*, 28(1), 333-361.
- Brodski, A., Paasch, G. F., Helbling, S., & Wibrat, M. (2015). The faces of predictive coding. *Journal of Neuroscience*, 35(24), 8997-9006.
- Brooks, J. A., & Freeman, J. B. (2018). Conceptual knowledge predicts the representational structure of facial emotion perception. *Nature human behaviour*, 2(8), 581-591.
- Brooks, J. A., Stoller, R. M., & Freeman, J. B. (2018). Stereotypes bias visual prototypes for sex and emotion categories. *Social Cognition*, 36(5), 481-493.
- Brown, E., & Perrett, D. I. (1993). What gives a face its gender? *Perception*, 22(7), 829-840.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British journal of psychology*, 77(3), 305-327.
- Bukach, C. M., Phillips, W. S., & Gauthier, I. (2010). Limits of generalization between categories and implications for theories of category specificity. *Attention, Perception, & Psychophysics*, 72(7), 1865-1874.
- Burton, A. M., Bruce, V., & Hancock, P. J. (1999). From pixels to people: A model of familiar face recognition. *Cognitive science*, 23(1), 1-31.
- Busey, T. A., & Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision research*, 45(4), 431-448.
- Canas-Bajo, T.,
 Whitney, D. (2020). Stimulus-specific individual differences in holistic perception of Mooney faces. *Frontiers in Psychology*, 2964.
- Canas-Bajo, T. & Whitney, D. (2022). Relative tuning of holistic face processing towards the fovea. *Vision Research*, 197.
- Cavanagh, P. (1991). What's up in top-down processing. *Representations of vision: Trends and tacit assumptions in vision research*, 295-304.
- Cerf, M., Frady, E. P., & Koch, C. (2009). Faces and text attract gaze independent of the task: Experimental data and computer model. *Journal of vision*, 9(12), 10-10.

- Cerf, M., Harel, J., Einhäuser, W., & Koch, C. (2008). Predicting human gaze using low-level saliency combined with face detection. *Advances in neural information processing systems*, 20, 1-7.
- Chin, M. D., Evans, K. K., Wolfe, J. M., Bowen, J., & Tanaka, J. W. (2018). Inversion effects in the expert classification of mammograms and faces. *Cognitive research: principles and implications*, 3(1), 1-9.
- Costela, F. M., & Woods, R. L. (2019). When watching video, many saccades are curved and deviate from a velocity profile model. *Frontiers in neuroscience*, 12, 960.
- Coutrot, A., & Guyader, N. (2014). How saliency, faces, and sound influence gaze in dynamic social scenes. *Journal of vision*, 14(8), 5-5.
- Crookes, K., & McKone, E. (2009). Early maturity of face recognition: No childhood development of holistic processing, novel face encoding, or face-space. *Cognition*, 111(2), 219-247.
- Crouzet, S. M., Kirchner H., & Thorpe, S. J. (2010). Fast saccades toward faces: face detection in just 100 ms. *Journal of vision*, 10(4), 16-16.
- Dakin, S. C., & Watt, R. J. (2009). Biological “bar codes” in human faces. *Journal of vision*, 9(4), 2-2.
- Dallenbach, K.M (1951). A Puzzle-picture with a New Principle of Concealment. *American journal of psychology, Department of psychology, University of Texas*.
- Damasio, A. R., Damasio, H., & Van Hoesen, G. W. (1982). Prosopagnosia: anatomic basis and behavioral mechanisms. *Neurology*, 32(4), 331-331.
- DeGutis, J., Mercado, R. J., Wilmer, J., & Rosenblatt, A. (2013). Individual differences in holistic processing predict the own-race advantage in recognition memory. *PLoS One*, 8(4), e58253.
- DeGutis, J., Wilmer, J., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition*, 126(1), 87-100.
- De Heering, A., & Rossion, B. (2008). Prolonged visual experience in adulthood modulates holistic face perception. *PloS one*, 3(5), e2317.
- Dobs, K., Isik, L., Pantazis, D., & Kanwisher, N. (2019). How face perception unfolds over time. *Nature communications*, 10(1), 1-10.
- Donnelly, N., & Davidoff, J. (1999). The mental representations of faces and houses: Issues concerning parts and wholes. *Visual cognition*, 6(3-4), 319-343.
- Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological and Personality Science*, 3(5), 562-571.
- Dotsch, R., Wigboldus, D. H., Langner, O., & van Knippenberg, A. (2008). Ethnic out-group faces are biased in the prejudiced mind. *Psychological science*, 19(10), 978-980.
- Dotsch, R. O. N., Wigboldus, D. H., & Van Knippenberg, A. D. (2013). Behavioral information biases the expected facial appearance of members of novel groups. *European Journal of Social Psychology*, 43(1), 116-125.
- Edwards, G., VanRullen, R., & Cavanagh, P. (2018). Decoding trans-saccadic memory. *Journal of Neuroscience*, 38(5), 1114-1123.

- Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.
- Éthier-Majcher, C., Joubert, S., & Gosselin, F. (2013). Reverse correlating trustworthy faces in young and older adults. *Frontiers in psychology*, 4, 592.
- Farah, M. J. (1991). Cognitive neuropsychology: Patterns of co-occurrence among the associative agnosias: Implications for visual object representation. *Cognitive Neuropsychology*, 8(1), 1-19.
- Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human perception and performance*, 21(3), 628.
- Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological review*, 105(3), 482.
- Farzin, F., Rivera, S. M., & Whitney, D. (2009). Holistic crowding of Mooney faces. *Journal of vision*, 9(6), 18-18.
- Farzin, F., Rivera, S. M., & Whitney, D. (2010). Spatial resolution of conscious visual perception in infants. *Psychological science*, 21(10), 1502-1509.
- Ferguson, K. T., Kulkofsky, S., Cashon, C. H., & Casasola, M. (2009). The development of specialized processing of own-race faces in infancy. *Infancy*, 14(3), 263-284.
- Fischer, J., & Whitney, D. (2011). Object-level visual information gets through the bottleneck of crowding. *Journal of Neurophysiology*, 106(3), 1389-1398.
- Foulsham, T., Cheng, J. T., Tracy, J. L., Henrich, J., & Kingstone, A. (2010). Gaze allocation in a dynamic situation: Effects of social status and speaking. *Cognition*, 117(3), 319-331.
- Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition*, 110(2), 160-170.
- Freeman, J. B., Rule, N. O., Adams Jr, R. B., & Ambady, N. (2010). The neural basis of categorical face perception: Graded representations of face gender in fusiform and orbitofrontal cortices. *Cerebral Cortex*, 20(6), 1314-1322.
- Ganel, T., & Goshen-Gottstein, Y. (2002). Perceptual integrality of sex and identity of faces: further evidence for the single-route hypothesis. *Journal of Experimental Psychology: Human Perception and Performance*, 28(4), 854.
- Gauthier, I. (2020). What we could learn about holistic face processing only from nonface objects. *Current Directions in Psychological Science*, 29(4), 419-425.
- George, N., Jemel, B., Fiori, N., Chaby, L., & Renault, B. (2005). Electrophysiological correlates of facial decision: insights from upright and upside-down Mooney-face perception. *Cognitive Brain Research*, 24(3), 663-673.
- Goffaux, V., Hault, B., Michel, C., Vuong, Q. C., & Rossion, B. (2005). The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. *Perception*, 34(1), 77-86.
- Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Signal but not noise changes with perceptual learning. *Nature*, 402(6758), 176-178.
- Gold, J. M., Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2000). Deriving behavioural receptive fields for visually completed contours. *Current Biology*, 10(11), 663-666.

- Gomez, J., Natu, V., Jeska, B., Barnett, M., & Grill-Spector, K. (2018). Development differentially sculpts receptive fields across early and high-level human visual cortex. *Nature communications*, 9(1), 1-12.
- Goodman, G. S., Sayfan, L., Lee, J. S., Sandhei, M., Walle-Olsen, A., Magnussen, S., ... & Arredondo, P. (2007). The development of memory for own-and other-race faces. *Journal of Experimental Child Psychology*, 98(4), 233-242.
- Goold, J. E., & Meng, M. (2016). Visual Search of Mooney Faces. *Frontiers in psychology*, 7, 155. doi:10.3389/fpsyg.2016.00155
- Gosselin, F., & Schyns, P. G. (2003). Superstitious perceptions reveal properties of internal representations. *Psychological science*, 14(5), 505-509.
- Gregory, R. L. (1970). *The intelligent eye*. McGraw-Hill.
- Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience*, 15(8), 536-548.
- Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Eccentricity bias as an organizing principle for human high-order object areas. *Neuron*, 34(3), 479-490.
- Hershler, O., Golan, T., Bentin, S., & Hochstein, S. (2010). The wide window of face detection. *Journal of Vision*, 10(10), 21-21.
- Hole, G. J., George, P. A., Eaves, K., & Rasek, A. (2002). Effects of geometric distortions on face-recognition performance. *Perception*, 31(10), 1221-1240.
- Hu, F., Hu, H., Xu, L., & Qin, J. (2013). The asymmetric distribution of informative face information during gender recognition. *Perceptual and Motor Skills*, 116(1), 106-117.
- Jack, R. E., Caldara, R., & Schyns, P. G. (2012). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology: General*, 141(1), 19.
- Jack, R. E., & Schyns, P. G. (2015). The human face as a dynamic tool for social communication. *Current Biology*, 25(14), R621-R634.
- Kalpadakis-Smith, A. V., Goffaux, V., & Greenwood, J. A. (2018). Crowding for faces is determined by visual (not holistic) similarity: Evidence from judgements of eye position. *Scientific reports*, 8(1), 1-14.
- Kamps, F. S., Hendrix, C. L., Brennan, P. A., & Dilks, D. D. (2020). Connectivity at the origins of domain specificity in the cortical face and place networks. *Proceedings of the National Academy of Sciences*, 117(11), 6163-6169.
- Kanwisher, N. (2001). Faces and places: of central (and peripheral) interest. *Nature neuroscience*, 4(5), 455-456.
- Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition*, 68(1), B1-B11.
- Karremans, J. C., Dotsch, R., & Corneille, O. (2011). Romantic relationship status biases memory of faces of attractive opposite-sex others: evidence from a reverse-correlation paradigm. *Cognition*, 121(3), 422-426.
- Kay, K. N., Weiner, K. S., & Grill-Spector, K. (2015). Attention reduces spatial uncertainty in human ventral temporal cortex. *Current Biology*, 25(5), 595-600.

- Ke, T. W., Stella, X. Y., & Whitney, D. (2017, September). “Mooney face classification and prediction by learning across tone”. In *2017 IEEE International Conference on Image Processing (ICIP)*; September 17, 2017; IEEE, 2025-2029.
- Kleiner, M., Brainard, D., & Pelli, D. (2007). What’s new in Psychtoolbox-3? *Perception*, *36*, 1-16.
- Kok, E. M., Van Geel, K., van Merriënboer, J. J., & Robben, S. G. (2017). What we do and do not know about teaching medical image interpretation. *Frontiers in psychology*, *8*, 309.
- Kovács, P., Knakker, B., Hermann, P., Kovács, G., & Vidnyánszky, Z. (2017). Face inversion reveals holistic processing of peripheral faces. *Cortex*, *97*, 81-95.
- Kreichman, O., Bonnef, Y. S., & Gilaie-Dotan, S. (2020). Investigating face and house discrimination at foveal to parafoveal locations reveals category-specific characteristics. *Scientific Reports*, *10*(1), 1-15.
- Kuefner, D., Cassia, V. M., Vescovo, E., & Picozzi, M. (2010). Natural experience acquired in adulthood enhances holistic processing of other-age faces. *Visual Cognition*, *18*(1), 11-25.
- Kwon, M., & Legge, G. E. (2011). Spatial-frequency cutoff requirements for pattern recognition in central and peripheral vision. *Vision research*, *51*(18), 1995-2007.
- Landi, S. M., & Freiwald, W. A. (2017). Two areas for familiar face recognition in the primate brain. *Science*, *357*(6351), 591-595.
- Lansing, C. R., & McConkie, G. W. (2003). Word identification and eye fixation locations in visual and visual-plus-auditory presentations of spoken sentences. *Perception & psychophysics*, *65*(4), 536-552.
- Latinus, M., & Taylor, M. J. (2005). Holistic processing of faces: learning effects with Mooney faces. *Journal of cognitive neuroscience*, *17*(8), 1316-1327.
- Latinus, M., & Taylor, M. J. (2006). Face processing stages: impact of difficulty and the separation of effects. *Brain research*, *1123*(1), 179-187.
- Laurence, S., Baker, K. A., Proietti, V. M., & Mondloch, C. J. (2021). What happens to our representation of identity as familiar faces age? Evidence from priming and identity aftereffects. *British Journal of Psychology*.
- Leo, I., & Simion, F. (2009). Newborns’ Mooney-Face Perception. *Infancy*, *14*(6), 641-653.
- Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature neuroscience*, *4*(5), 533-539.
- Liu, P., Montaser-Kouhsari, L., & Xu, H. (2014). Effects of face feature and contour crowding in facial expression adaptation. *Vision research*, *105*, 189-198.
- Logothetis, N. K., Pauls, J., Bülthoff, H. H., & Poggio, T. (1994). View-dependent object recognition by monkeys. *Current biology*, *4*(5), 401-414.
- Louie, E. G., Bressler, D. W., & Whitney, D. (2007). Holistic crowding: Selective interference between configural representations of faces in crowded scenes. *Journal of Vision*, *7*(2), 24-24.
- Malach, R., Levy, I., & Hasson, U. (2002). The topography of high-order human object areas. *Trends in cognitive sciences*, *6*(4), 176-184.

- Manassi, M., & Whitney, D. (2018). Multi-level crowding and the paradox of object recognition in clutter. *Current Biology*, 28(3), 127-133.
- Mangini, M. C., & Biederman, I. (2004). Making the ineffable explicit: Estimating the information employed for face classifications. *Cognitive Science*, 28(2), 209-226.
- Marat, S., Rahman, A., Pellerin, D., Guyader, N., & Houzet, D. (2013). Improving visual saliency by adding ‘face feature map’ and ‘center bias’. *Cognitive Computation*, 5(1), 63-75.
- Mareschal, I., Dakin, S. C., & Bex, P. J. (2006). Dynamic properties of orientation discrimination assessed by using classification images. *Proceedings of the National Academy of Sciences*, 103(13), 5131-5136.
- Martelli, M., Majaaj, N. J., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision*, 5(1), 6-6.
- Martin, J. G., Davis, C. E., Riesenhuber, M., & Thorpe, S. J. (2018). Zapping 500 faces in less than 100 seconds: Evidence for extremely fast and sustained continuous visual search. *Scientific reports*, 8(1), 1-12.
- Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in cognitive sciences*, 6(6), 255-260.
- McCaffery, J. M., Robertson, D. J., Young, A. W., & Burton, A. M. (2018). Individual differences in face identity processing. *Cognitive research: principles and implications*, 3(1), 21.
- McKone, E. (2004). Isolating the special component of face recognition: peripheral identification and a Mooney face. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(1), 181.
- McKone, E. (2008). Configural processing and face viewpoint. *Journal of Experimental Psychology: Human Perception and Performance*, 34(2), 310.
- McKone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing. *Psychonomic bulletin & review*, 16(5), 778-797.
- Michel, C., Rossion, B., Han, J., Chung, C. S., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one’s own race. *Psychological Science*, 17(7), 608-615.
- Mooney, C. M. (1957). Age in the development of closure ability in children. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 11(4), 219.
- Moore, C., & Cavanagh, P. (1998). Recovery of 3D volume from 2-tone images of novel objects. *Cognition*, 67(1-2), 45-71.
- Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of cognitive neuroscience*, 9(5), 555-604.
- Murray, R. F. (2011). Classification images: A review. *Journal of vision*, 11(5), 2-2.
- Neri, P., & Heeger, D. J. (2002). Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nature neuroscience*, 5(8), 812-816.
- Nestor, A., & Tarr, M. J. (2008). The segmental structure of faces and its use in gender recognition. *Journal of Vision*, 8(7), 7-7: 1-12.

- Nichols, D. F., Betts, L. R., & Wilson, H. R. (2010). Decoding of faces and face components in face-sensitive human visual cortex. *Frontiers in Psychology*, 1, 28.
- Orrù, G., Marcialis, G. L., & Roli, F. (2020). A novel classification-selection approach for the self-updating of template-based face recognition systems. *Pattern Recognition*, 100, 107121.
- Otsuka, Y., Hill, H. C., Kanazawa, S., Yamaguchi, M. K., & Spehar, B. (2012). Perception of Mooney faces by young infants: The role of local feature visibility, contrast polarity, and motion. *Journal of Experimental Child Psychology*, 111(2), 164-179.
- Pascalis, O., & Kelly, D. J. (2009). The origins of face processing in humans: Phylogeny and ontogeny. *Perspectives on Psychological Science*, 4(2), 200-209.
- Phillips, P. J., & O'toole, A. J. (2014). Comparison of human and computer performance across face recognition experiments. *Image and Vision Computing*, 32(1), 74-85.
- Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature*, 343(6255), 263-266.
- Poltoratski, S., Kay, K. & Grill-Spector, K. (2019) Using population receptive field models to elucidate spatial integration in high-level visual cortex. *Conference on Cognitive Computational Neuroscience*, 13-16 September 2019, Berlin, Germany.
- Ramon, M., Vizioli, L., Liu-Shuang, J., & Rossion, B. (2015). Neural microgenesis of personally familiar face recognition. *Proceedings of the National Academy of Sciences*, 112(35), E4835-E4844.
- Ratner, K. G., Dotsch, R., Wigboldus, D. H., van Knippenberg, A., & Amodio, D. M. (2014). Visualizing minimal ingroup and outgroup faces: Implications for impressions, attitudes, and behavior. *Journal of personality and social psychology*, 106(6), 897.
- Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychological science*, 22(4), 464-471.
- Ringach, D., & Shapley, R. (2004). Reverse correlation in neurophysiology. *Cognitive Science*, 28(2), 147-166.
- Rivest, J., Moscovitch, M., & Black, S. (2009). A comparative case study of face recognition: the contribution of configural and part-based recognition systems, and their interaction. *Neuropsychologia*, 47(13), 2798-2811.
- Ro, T., Russell, C., & Lavie, N. (2001). Changing faces: A detection advantage in the flicker paradigm. *Psychological science*, 12(1), 94-99.
- Rock, I., & DiVita, J. (1987). A case of viewer-centered object perception. *Cognitive psychology*, 19(2), 280-293.
- Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta psychologica*, 128(2), 274-289.
- Rossion, B. (2009). Distinguishing the cause and consequence of face inversion: The perceptual field hypothesis. *Acta psychologica*, 132(3), 300-312.
- Rossion, B., & Boremanse, A. (2008). Nonlinear relationship between holistic processing of individual faces and picture-plane rotation: Evidence from the face composite illusion. *Journal of vision*, 8(4), 3-3.

- Russell, R., Duchaine, B., & Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic bulletin & review*, 16(2), 252-257.
- Sayres, R., & Grill-Spector, K. (2008). Relating retinotopic and object-selective responses in human lateral occipital cortex. *Journal of neurophysiology*, 100(1), 249-267.
- Schwartz, J. L., Berthommier, F., & Savariaux, C. (2004). Seeing to hear better: evidence for early audio-visual interactions in speech identification. *Cognition*, 93(2), B69-B78.
- Schwarzlose, R. F., Swisher, J. D., Dang, S., & Kanwisher, N. (2008). The distribution of category and location information across object-selective regions in human visual cortex. *Proceedings of the National Academy of Sciences*, 105(11), 4447-4452.
- Schwiedrzik, C. M., Melloni, L., & Schurger, A. (2018). Mooney face stimuli for visual perception research. *PloS one*, 13(7), e0200106.
- Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, 13(5), 402-409.
- Sekuler, R. W., & Abrams, M. (1968). Visual sameness: a choice time analysis of pattern recognition processes. *Journal of Experimental Psychology*, 77(2), 232.
- Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, 14(5), 391-396.
- Sekuler, A. B., Gold, J. M., Murray, R. F., & Bennett, P. (2000). Visual completion of partly occluded objects: Insights from behavioral studies. *Neuroophthalmology*, 23, 165-168.
- Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology*, 75(2), 221-242.
- Sergent, J., & Signoret, J. L. (1992). Varieties of functional deficits in prosopagnosia. *Cerebral cortex*, 2(5), 375-388.
- Sheridan, H., & Reingold, E. M. (2017). The holistic processing account of visual expertise in medical image perception: A review. *Frontiers in psychology*, 8, 1620.
- Silson, E. H., Groen, I. I., Kravitz, D. J., & Baker, C. I. (2016). Evaluating the correspondence between face-, scene-, and object-selectivity and retinotopic organization within lateral occipitotemporal cortex. *Journal of vision*, 16(6), 14-14.
- Simpson, W. E., & Crandall, S. J. (1972). The perception of smiles. *Psychonomic Science*, 29(4), 197-200.
- Smith, E. E., & Nielsen, G. D. (1970). Representations and retrieval processes in short-term memory: Recognition and recall of faces. *Journal of Experimental Psychology*, 85(3), 397.
- Solomon, J. A. (2002). Noise reveals visual mechanisms of detection and discrimination. *Journal of vision*, 2(1), 7-7.
- Steeves, J. K., Culham, J. C., Duchaine, B. C., Pratesi, C. C., Valyear, K. F., Schindler, I., ... & Goodale, M. A. (2006). The fusiform face area is not sufficient for face recognition: evidence from a patient with dense prosopagnosia and no occipital face area. *Neuropsychologia*, 44(4), 594-609.

- Susilo, T., Crookes, K., McKone, E., & Turner, H. (2009). The composite task reveals stronger holistic processing in children than adults for child faces. *PLoS One*, 4(7), e6460.
- Tarr, M. J. (1995). Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin & Review*, 2(1), 55-82.
- Tarr, M. J., & Bülthoff, H. H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views? *Comment on Biederman and Gerhardstein (1993)*.
- Taubert, J., Apthorp, D., Aagten-Murphy, D., & Alais, D. (2011). The role of holistic processing in face perception: Evidence from the face inversion effect. *Vision research*, 51(11), 1273-1278.
- Theeuwes, J., & Van der Stigchel, S. (2006). Faces capture attention: Evidence from inhibition of return. *Visual Cognition*, 13(6), 657-665.
- Todorov, A., Dotsch, R., Wigboldus, D. H., & Said, C. P. (2011). Data-driven methods for modeling social perception. *Social and Personality Psychology Compass*, 5(10), 775-791.
- Treviño, M., Birdsong, G., Carrigan, A., Choyke, P., Drew, T., Eckstein, M., ... Wolfe, J. M. (2022). Advancing Research on Medical Image Perception by Strengthening Multi-disciplinary Collaboration. *JNCI Cancer Spectrum*, 6(1), pkab099.
- Turk, M. A., & Pentland, A. P. (1991, January). Face recognition using eigenfaces. In *Proceedings. 1991 IEEE computer society conference on computer vision and pattern recognition* (pp. 586-587). IEEE Computer Society.
- Umiltà, C., Simion, F., & Valenza, E. (1996). Newborn's preference for faces. *European Psychologist*, 1(3), 200-205.
- Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology Section A*, 43(2), 161-204.
- Verhallen, R. J., Bosten, J. M., Goodbourn, P. T., Bargary, G., Lawrance-Owen, A. J., & Mollon, J. D. (2014). An online version of the Mooney Face Test: phenotypic and genetic associations. *Neuropsychologia*, 63, 19-25.
- Verhallen, R. J., & Mollon, J. D. (2016). A new Mooney test. *Behavior research methods*, 48(4), 1546-1559.
- Victor, J. D. (2005). Analyzing receptive fields, classification images and functional images: challenges with opportunities for synergy. *Nature neuroscience*, 8(12), 1651-1656.
- Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International journal of computer vision*, 57(2), 137-154.
- Visconti di Oleggio Castello, M., Halchenko, Y. O., Guntupalli, J. S., Gors, J. D., & Gobbini, M. I. (2017). The neural representation of personally familiar and unfamiliar faces in the distributed system for face perception. *Scientific reports*, 7(1), 1-14.
- Waite, S., Grigorian, A., Alexander, R. G., Macknik, S. L., Carrasco, M., Heeger, D. J., & Martinez-Conde, S. (2019). Analysis of perceptual expertise in radiology—Current knowledge and a new perspective. *Frontiers in human neuroscience*, 13, 213.

- Wang, R., Li, J., Fang, H., Tian, M., & Liu, J. (2012). Individual differences in holistic processing predict face recognition ability. *Psychological science*, 23(2), 169-177.
- Watson, A. B. (1998). Multi-category classification: Template models and classification images. *Investigative Ophthalmology and Visual Science*, 39(ARVO Suppl. 4), S912.
- Williams, M. A., Baker, C. I., De Beeck, H. P. O., Shim, W. M., Dang, S., Triantafyllou, C., & Kanwisher, N. (2008). Feedback of visual object information to foveal retinotopic cortex. *Nature neuroscience*, 11(12), 1439.
- Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., ... & Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of sciences*, 107(11), 5238-5241.
- Wolfe, B. A., & Whitney, D. (2014). Facilitating recognition of crowded faces with presaccadic attention. *Frontiers in Human Neuroscience*, 8, 103.
- Wolfe, B. A., & Whitney, D. (2015). Saccadic remapping of object-selective information. *Attention, Perception, & Psychophysics*, 77(7), 2260-2269.
- Wong, A. C. N., Palmeri, T. J., & Gauthier, I. (2009). Conditions for facelike expertise with objects: Becoming a Ziggerin expert—but which type?. *Psychological Science*, 20(9), 1108-1117.
- Xia, Y., Manassi, M., Nakayama, K., Zipser, K., & Whitney, D. (2020). Visual crowding in driving. *Journal of vision*, 20(6), 1-1.
- Yamaguchi, M. K., Hirukawa, T., & Kanazawa, S. (2013). Judgment of gender through facial parts. *Perception*, 42(11), 1253-1265.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of experimental psychology*, 81(1), 141.
- Yokoyama, T., Noguchi, Y., Tachibana, R., Mukaida, S., & Kita, S. (2014). A critical role of holistic processing in face gender perception. *Frontiers in human neuroscience*, 8, 477.
- Young, A. W., & Bruce, V. (2011). Understanding person perception. *British journal of psychology*, 102(4), 959-974.
- Yovel, G., Wilmer, J. B., & Duchaine, B. (2014). What can individual differences reveal about face processing? *Frontiers in human neuroscience*, 8, 562.
- Yue, X., Cassidy, B. S., Devaney, K. J., Holt, D. J., & Tootell, R. B. (2011). Lower-level stimulus features strongly influence responses in the fusiform face area. *Cerebral Cortex*, 21(1), 35-47.
- Zhao, M., & Hayward, W. G. (2010). Holistic processing underlies gender judgments of faces. *Attention, Perception, & Psychophysics*, 72(3), 591-596.
- Zhao, M., & Hayward, W. G. (2013). Integrative processing of invariant aspects of faces: Effect of gender and race processing on identity analysis. *Journal of Vision*, 13(1), 15-15.
- Zhou, G., Cheng, Z., Zhang, X., & Wong, A. C. N. (2012). Smaller holistic processing of faces associated with face drawing experience. *Psychonomic bulletin & review*, 19(2), 157-162.