Ensemble statistics of faces

Ying, Haojiang

2019

Ying, H. (2019). Ensemble statistics of faces. Doctoral thesis, Nanyang Technological University, Singapore.

https://hdl.handle.net/10356/89893

https://doi.org/10.32657/10220/47740

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

ENSEMBLE STATISTICS OF FACES

HAOJIANG YING

SCHOOL OF SOCIAL SCIENCES

2019

ENSEMBLE STATISTICS OF FACES HAOJIANG YING 2019

HAOJIANG YING

Ensemble Statistics of Faces

HAOJIANG YING

School of Social Sciences

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2019

Statement of Originality

I certify that all work submitted for this thesis is my original work. I declare that no other person's work has been used without due acknowledgement. Except where it is clearly stated that I have used some of this material elsewhere, this work has not been presented by me for assessment in any other institution or University. I certify that the data collected for this project are authentic and the investigations were conducted in accordance with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

2019/02/11 Huang Ying

Date [Student's Name Here]

Supervisor Declaration Statement

I have reviewed the content of this thesis and to the best of my knowledge, it does not contain plagiarised materials. The presentation style is also consistent with what is expected of the degree awarded. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accordance with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

. . . .11/02/2019 . . .   . . Xu Hong .

Date   [Supervisor Name Here]

Authorship Attribution Statement

This thesis contains material from 2 papers published in the following peer-reviewed journals where I was the first author.

Chapter II is published as: Ying, H., & Xu, H. (2017). Adaptation reveals that facial expression averaging occurs during rapid serial presentation. *Journal of Vision, 17*, 15. doi:10.1167/17.1.15

The contributions of the co-authors are as follows:

- I and Asst/Prof Xu provided the initial project direction. I prepared the manuscript drafts. The manuscript was revised by Asst/Prof Xu.
- I co-designed the study with Asst/Prof Xu and performed all the laboratory work at the School of Social Sciences. I also analyzed the data.

Chapter IV is published as: Ying, H., Burns, E. J., Lin, X., & Xu, H. (In print). Ensemble statistics shapes face adaptation and the cheerleader effect. *Journal of Experimental Psychology: General.*

The contributions of the co-authors are as follows:

- I, Dr. Burns, Ms Lin, and Asst/Prof Xu developed the study concept and contributed to the study design.
- I performed the testing and data collection, with help from Ms Lin, at the School of Social Sciences.
- I performed the data analyses and interpretation under the supervision of Asst/Prof Xu.
- I drafted the article, Dr. Burns and Asst/Prof Xu provided critical revisions.

2019/02/11 .... | ....
Date | [Student's Name Here]

Acknowledgements

I would like to take this opportunity to express my gratitude to all of the people who helped me throughout my four-year studying here at Nanyang Technological University.

Firstly, I would like to thank my dear advisor, Prof Hong Xu. I deeply appreciate the valuable time she spent, and the constructive suggestions she offered. I learnt to be a researcher with her guidance. It was an honor for me to learn from her. Secondly, I would like to thank my TAC members and the professors in our department for their kind suggestions and helpful comments in research. Thirdly, I would like to thank Prof David Perrett from University of St Andrews and Prof Lisa DeBruine from University of Glasgow for their support in research.

I sincerely thank all my lab members, colleagues and friends for their help and time in these four years. Specifically, I would like to thank Bo, Daisy, Edwin, FeiTong, Ka Lon, RuYuan, JieXin, MengDan, Yang, Paul, Nadine, and many other friends for their continued support and the memories we shared.

Finally, I would like to thank my family members for their unconditioned love and support.

Table of Contents

Abstract

Ensemble statistics describes the ability of the visual system to summarize the information provided by a group of objects. This thesis refines the theoretical framework of ensemble statistics of faces throughout four studies.

Study 1 investigated involuntary ensemble statistics of facial expressions during rapid serial visual presentation (RSVP). Prolonged exposure to RSVP of faces led to significant facial expression adaptation aftereffects, which could be explained by ensemble statistics. Further testing clarified that this representation was only modulated by the mean information in that stream. In summary, Study 1 examined the involuntary ensemble coding of facial expressions and hinted at the potential mechanism behind the formation of face space.

Comparing adaptation aftereffects, Study 2 showed that temporal and spatial ensemble statistics of faces arise from distinct mechanisms that produce qualitatively

different perceptual outcomes. The visual system extracts the low-level ‘computational’ average from faces when faces are presented individually across time. However, the spatial ensemble statistics summarizes the higher-level gist. Study 2, for the first time, showed there are distinctive mechanisms for ensemble coding of the same facial charateristics.

Studying facial attractiveness adaptation and the ‘cheerleader effect’ (i.e. faces are perceived as more attractive when surrounded by others rather than being alone), Study 3 linked these two important phenomena in facial attractiveness with ensemble statistics. The mean attractiveness of the crowd (determined by ensemble statistics) biased the perceived attractiveness of the subsequently viewed face (the adaptation aftereffect). Similarly, the levels of attractiveness in a simultaneously presented crowd could also affect a target face (the ‘cheerleader effect’). Both past and present experience (determined by ensemble statistics) therefore impact the perception of facial attractiveness. Study 3 showed that the ensemble coding is ubiquitous in face perception and shapes two important phenomena in facial attractiveness. Also, the findings suggested how external factors, the previous and present exposures to faces, affect the attractiveness perception.

Converging evidence from Study 4 emphasized the role of attention in both explicit and implicit ensemble statistics of facial expressions, and also suggested that the ensemble statistics is a weighted average of the visual input rather than the simple average. This study further clarified the averaging mechanism of ensemble coding and unveiled the relationship between ensemble coding of the face and attention.

Taken together, the thesis examined the mechanisms involved in high-level ensemble statistics and the relationship among ensemble statistics, face perception, and attention. The results may shed light on a more comprehensive understanding of face perception and visual processing.

List of Figures

Publications

Chapter II:

Ying, H., & Xu, H. (2017). Adaptation reveals that facial expression averaging occurs during rapid serial presentation. *Journal of Vision, 17,* 15. doi:10.1167/17.1.15

All the figures of Chapter II are adapted from this paper with permission.

Chapter III:

Ying, H., Burns, E. J., Choo, A., & Xu, H. (Under Review). Temporal and spatial ensemble statistics are formed by distinct neural calculations. *Cognition.*

Chapter IV:

Ying, H., Burns, E. J., Lin, X., & Xu, H. (In print). Ensemble statistics shapes face adaptation and the cheerleader effect. *Journal of Experimental Psychology: General.*

Chapter V:

Ying, H., & Xu, H. (In Preparation). Attention Modulates the Ensemble Coding of Facial Expressions.

Chapter I: Introduction and Literature Review

1.1. Overview

Equipped with distributed neural systems for face perception, the human visual system is capable of obtaining valuable information from another individual's face quickly and effortlessly (Bruce & Young 1986; Calder & Young, 2005; Haxby, Hoffman, & Gobbini, 2000; Hoffman & Haxby, 2000; Leopold & Rhodes, 2010; Little, Jones, & DeBruine, 2011; Willis & Todorov, 2006; Young & Bruce, 2011). Within the repertoire of face perception, facial expression and facial attractiveness, due to their social and evolutionary significance, are vital for social communications. Facial expression, as a changeable aspect of faces, reflects not only the current mood, but also an individual's possible intentions (Elias, Dyer, & Sweeny, 2017; Jack & Schyns, 2017). On the other hand, facial attractiveness, as an invariant aspect of faces, is linked with other personality attributes and has a huge impact on mate choices as well as decisions about social exchange (Little, Jones, & DeBruine, 2011; Rhodes, 2006; Thornhill & Gangestad, 1999; Willis & Todorov, 2006).

To better understand the perception of facial expressions and attractiveness, it is necessary to clarify how faces are perceived in the group context, because a face seldom appears alone in the real world (Elias et al., 2017). However, the current understanding of facial expressions and attractiveness is mainly based on experiments testing the perception of a single face. How do we interpret groups of faces we see every day? When individuals are exposed to a large amount of information, perception is narrowed down to a finite number of objects, possibly due to the limited capacity of visual processing and working memory (Brady & Alvarez, 2015; Cohen, Dennett, & Kanwisher, 2016; Haberman & Whitney, 2012; Nieuwenstein & Potter, 2006). Therefore, there is a vast discrepancy between the perceived richness of the scene and our limited perceptual ability (Alvarez, 2011; Whitney & Leib, 2017).

Alternatively, the human visual system may process such a mass of information via ensemble statistics: obtaining the "gist" of the visual input (Alvarez & Oliva, 2009; Haberman & Whitney, 2012). Such ensemble coding has been observed in low-level features like orientation and direction, and high-level features like facial expression, identity and gaze viewpoint (Alvarez, 2011; Alvarez & Oliva, 2008, 2009; Ariely, 2001; Elias et al., 2017; Haberman, Brady, & Alvarez, 2015; Haberman, Harp, & Whitney, 2009; Haberman & Whitney, 2007, 2012; Sweeny & Whitney, 2014; Leib et al., 2014; Qiu, Robertson, & Whitney, 2014). For instance, Haberman and Whitney (2007) showed that subjects could correctly report the mean expression of a crowd at the expense of perception of each component face. Studies in statistical averaging have concentrated mainly on spatial ensembles (Alvarez & Oliva, 2008; Ariely, 2001; Chong & Treisman, 2003; Haberman et al., 2015; Leib et al., 2014), but relatively less on temporal ensembles (e.g.., Haberman, Harp, & Whitney, 2009). Considering the fact that faces are always temporally adjacent to other faces, Chapter

II (Study 1) aims to study whether and how temporal ensemble coding explains the perception of facial expressions in Rapid Serial Visual Presentation (RSVP).

The detailed mechanisms of ensemble statistics of faces are not clear at the moment. Despite the fact that researchers have described ensemble statistics as extracting the gist, it is still unclear what the gist actually represents in face perception (Whitney & Leib, 2017). For example, does ensemble coding extract a general representation, whereby a high-level judgment is made for each of the faces before summarizing the overall gist from these judgments? Or, does the brain compute the average of low-level visual properties of each face first and then judge the group average? Due to the limitation of the facial characteristics employed, previous studies were unable to clarify which hypothesis is correct. For instances, facial emotion predicts the same outcome for the two competing hypotheses (Figure 1.1A). Here, Chapter III (Study 2) takes advantage of the fact that an averaged face created from the low-level properties of a group of faces is more attractive than the gist of the underlying faces in the group (DeBruine, Jones, Unger, Little, & Feinberg, 2007). By measuring the perceived facial attractiveness of temporally and spatially presented faces using an adaptation paradigm, Study 2 investigated for the first time what the possible computational mechanisms of ensemble statistics of faces are.

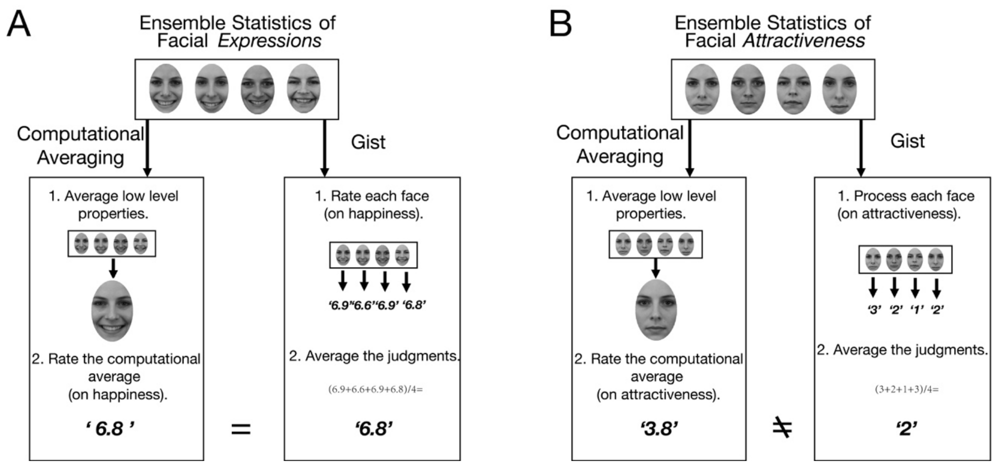

Figure 1.1. The 'Computational Average' and 'Gist' hypotheses (the demonstrated faces are AF01NES, AF05NES, AF06NES and AF07NES from KDEF database). All the ratings are hypothetical and only for illustration purposes. (A) Ensemble coding for facial expressions: the 'Computational Averaging' and 'Gist' hypotheses predict the same perceptual outcome for emotion; i.e., happy intensity rating of 6.8. (B) Ensemble coding of facial attractiveness: the computational average face is more attractive than the mean attractiveness of its individual component faces, with the computational average not equal to the gist (i.e., attractiveness rating of 3.8 versus 2).

Recent evidence suggests that ensemble coding is actively involved in many aspects of visual processing (Whitney & Leib, 2017). For instance, Fischer and Whitney (2011) have suggested that ensemble statistics underpin the visual crowding (Whitney & Levi, 2011). Therefore, it is reasonable to examine whether ensemble coding is involved in face perception, especially in facial attractiveness. Because of the distinctive neural systems for face perception, one should not assume that the ensemble statistics shape facial attractiveness in the same way as they shape facial expression. In Chapter IV (Study 3), four experiments were conducted to link ensemble statistics with essential phenomena in facial attractiveness.

What is the relationship between attention and ensemble statistics of faces? So far, different researchers have found mixed results. It has been shown that ensemble coding can occur with reduced attention (Alvarez & Oliva, 2009). However, it might not be correct to assume that ensemble coding is immune to attention control. Recent evidence has revealed that attention modulates the ensemble coding of size (Chong & Treisman, 2005; de Fockert & Marchant, 2008; Li & Yeh, 2017). Furthermore, Elias and colleagues (2017) showed that the averaging of facial emotions could be disrupted when attention was diverted by a secondary task. To clarify the relationship between attention and ensemble coding of faces, Chapter V (Study 4) tested the extent to which exogenous and endogenous attentional cues modulate the ensemble statistics of facial expressions.

In summary, this thesis examines the spatial and temporal ensemble statistics of facial expressions and attractiveness and investigates the mechanisms of ensemble statistics as well as face perception. Four studies, including thirteen experiments, were conducted and detailed in the following four chapters (illustrated in Fig 1.2).

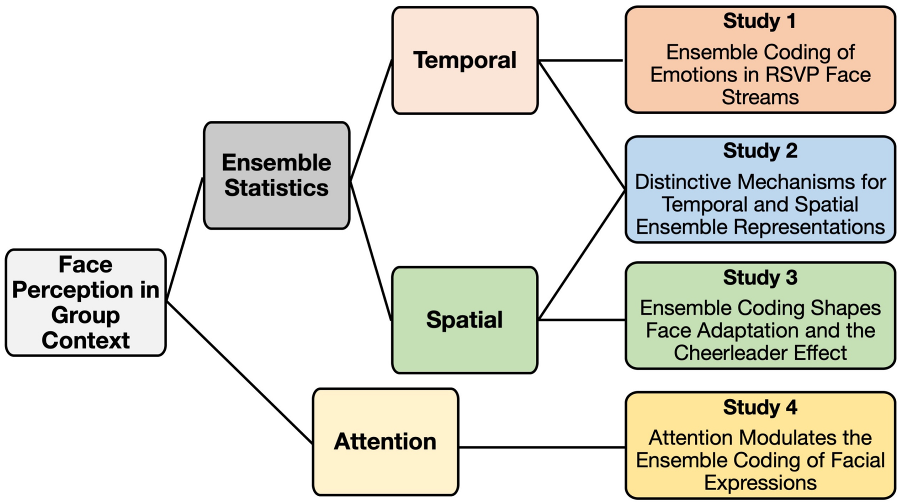

Figure 1.2. Illustration of the research framework.

Chapter II (Study 1) investigated the processing of facial expression during rapid serial visual presentation (RSVP). After passively viewing streams of faces in rapid succession (42.5 Hz, each face for 23.5 ms only), subject's perception of the subsequently viewed test face was biased, revealing a facial expression adaptation aftereffect (Experiment 1.1). The aftereffect could be explained by ensemble statistics (Experiment 1.2): the RSVP of faces and the computer-generated averaged faces of the RSVP stream evoked equivalent and correlated adaptation aftereffects. Controls in emotion variation, the temporal frequency of the stream, and the identity of the testing stimuli further clarified that the representation of the stream is only subject to the mean of that stream (Experiment 1.3 & 1.4). These findings suggested that ensemble statistics are involuntarily used to interpret rapidly changing faces.

Chapter III (Study 2) aimed to clarify the computational mechanism of ensemble statistics. Taking advantage of the fact that the average of unattractive faces is more attractive than a given component, Study 2 differentiated the 'computational averaging' (average the low-level features and then judge) and gist (perceive each face and then average the judgments) hypotheses. Results from Experiment 2.1 & 2.2 showed that the visual system extracts the low-level 'computational average' from faces when faces are presented individually across time. However, the spatial ensemble statistics summarizes the higher level gist (Experiment 2.3). Temporal and spatial ensemble statistics of faces, therefore, arise from distinct mechanisms that produce qualitatively different perceptual outcomes.

Chapter IV (Study 3) explored the relationship between ensemble statistics and two important phenomena in facial attractiveness, the adaptation aftereffect and the "cheerleader effect", and the degree to which these two phenomena are associated. The data from Experiment 3.1 & 3.2 revealed that the mean attractiveness of the crowd biases the perceived attractiveness of subsequently viewed face, revealing the facial attractiveness aftereffect. Experiment 3.3 & 3.4 further showed that the ensemble statistics of facial attractiveness also affect the "cheerleader effect": the more unattractive the crowding faces are, the more attractive the central face is perceived. The findings from the study revealed that both past and present experience (determined by ensemble statistics) impact the perception of facial attractiveness.

Chapter V (Study 4) investigated the impact of attention on the ensemble statistics of a crowd's expressions. The results showed that the reported mean emotion of the same crowd is heavily influenced by the orientation of the exogenous cues (Experiment 4.1): the cued face has a significantly larger weight in ensemble statistics. Using endogenous cues and the adaptation paradigm, Experiment 4.2

replicated the finding. Thus, converging evidence from this study emphasized the role of attention in the ensemble statistics of facial expressions, and also suggested that the ensemble statistics stem from a weighted average of the visual input rather than simply from the arithmetic mean.

The studies described in this thesis have examined the mechanisms of ensemble statistics of faces, linked the ensemble perception with important phenomena in face perception, and showed the ubiquity of ensemble statistics in face perception. The findings of the thesis could refine the theoretical frameworks of face perception and ensemble statistics. Moreover, this thesis directly targets the core concept of ensemble statistics: the definition and the formation of ‘averaging’ (i.e., ensemble representation). To the best of our knowledge, this thesis offers the first systematical analysis of the behavioral mechanisms of ensemble statistics of faces.

1.2. Literature Review

1.2.1 Functional and Neural models for Face perception

Our visual system has an impressive repertoire of face processing capabilties: individuals can derive identity, facial expression emotion, eye gaze direction, attractiveness, trustworthiness, kinship, age, sex dimorphism, and many other aspects by using slight physical differences among faces (Andrews & Ewbank, 2004; Bruce & Young, 1986; Buckingham et al., 2006; Welling, Conway, Tiddeman, & Jones, 2006; Burton, Jeffery, Calder, & Rhodes, 2014; Little et al., 2011; Maloney & Dal Martello, 2006; Oosterhof & Todorov, 2008; Rhodes & Jeffery, 2006; Rossion, 2014; Schyns, Petro, & Smith, 2007; Sutherland et al., 2013; Towler, Burt, & Young, 2013; Young & Bruce, 2011). Based on the converging evidence from psychophysics and neural imaging studies, there is a framework presented by Bruce and Young (1986) and more recently studies in the field (Calder & Young, 2005; Haxby et al., 2000;

Hoffman & Haxby, 2000; Young & Bruce, 2011). Using this model (briefly illustrated in Figure 1.3), researchers categorized face perception into two different domains based upon the nature of facial information: (1) "changeable aspects" of faces, such as facial expressions; and (2) "invariant aspects" of faces, such as facial attractiveness.

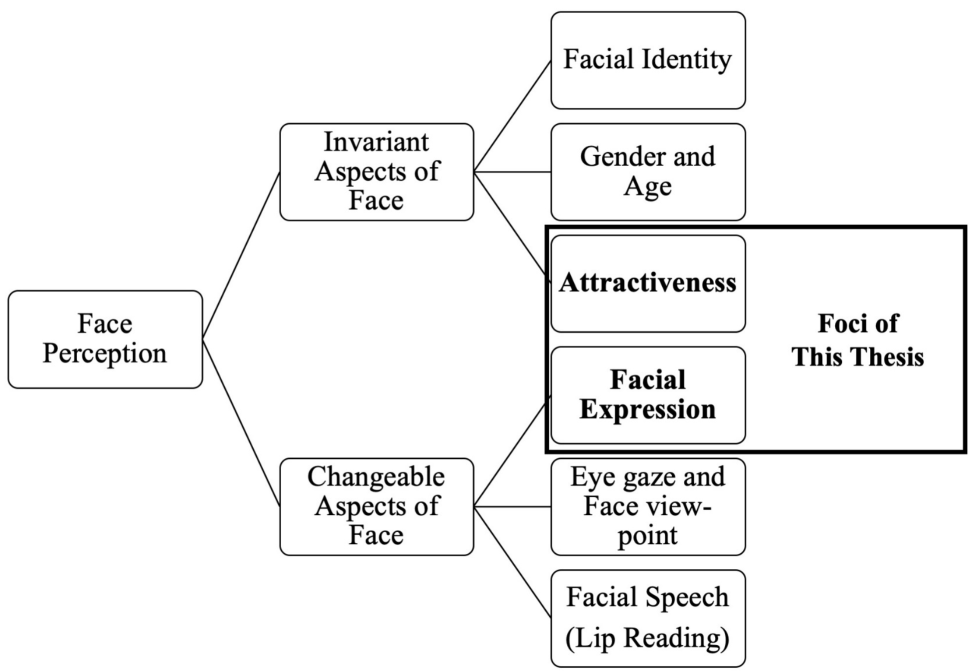

Figure 1.3. The framework of face perception. Based on the framework suggested by Bruce & Young, (1986) face perception can be categorized into two different domains by the nature of facial information: invariant aspects of the face (e.g. attractiveness), and changeable aspects of the face (e.g. facial expression). On the other

hand, both facial expression and facial attractiveness are vital for social communication and are the foci of this thesis.

Furthermore, Haxby, Hoffman, and Gobbini (2002), combining findings from neuroscience, have suggested that there are sophisticated circuits of distributed neural systems activating fine-grained face perception (Behrman & Plaut, 2013; Hoffman & Haxby, 2000; Haxby & Gobbini, 2011). In this model (Figure 1.4), there is a core system with specialized (face-dedicated) structures in the occipital and temporal lobes coding facial information; and there is also an extended system consisting of other brain structures involved in other cognitive functions aiding face perception.

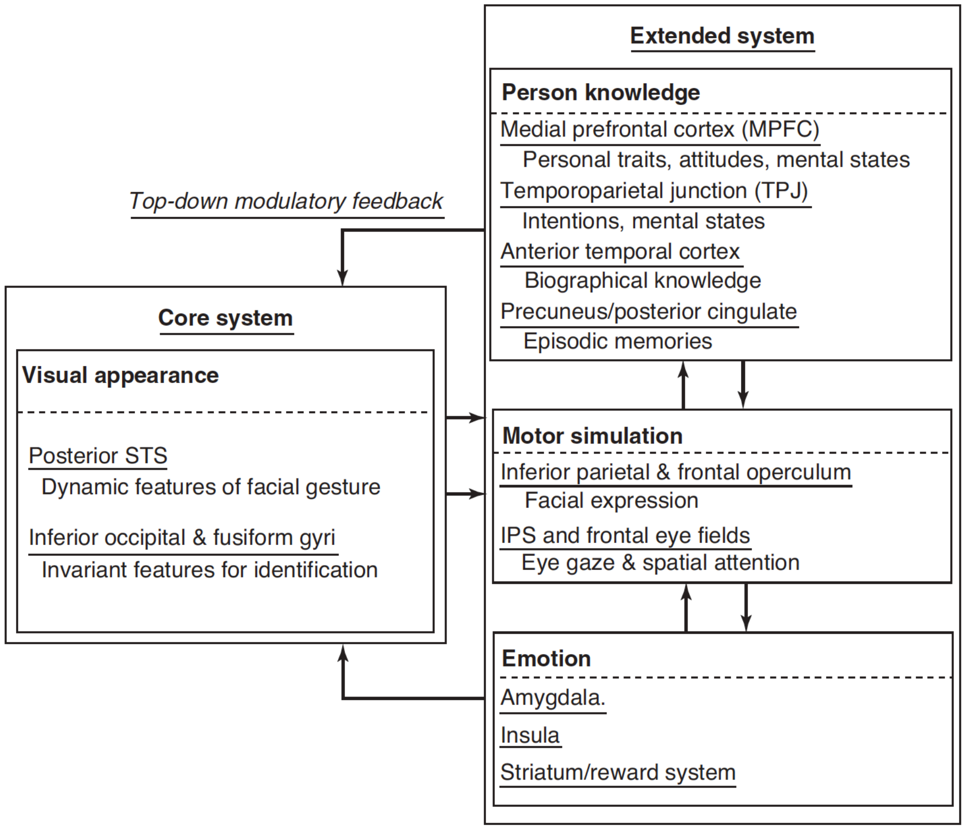

Figure 1.4. The model of distributed neural systems for face perception. Reprinted (with permission) from Haxby & Gobbini (2011).

The core system analyzes the facial information of both variant aspects and invariant aspects (Calder & Young, 2005). In the core system, the superior temporal sulcus (STS) is involved in the network for the perception of facial expression and gaze (Hoffman & Haxby, 2000; Perrett, Hietanen, Oram, Benson, & Rolls, 1992; Todorov, Gobbini, Evans, & Haxby, 2007; Vuilleumier, Armony, Driver, & Dolan, 2001), whereas the lateral fusiform gyrus, namely the fusiform face area (FFA; Kanwisher, McDermott, & Chun, 1997; see also Gauthier & Tarr, 1997), is in charge of the unchangeable aspects, such as facial identity perception (Andrews & Ewbank 2004; Liu, Harris, & Kanwisher, 2010). In contrast, the extended system helps to complete the face perception. For instance, Gobbini and colleagues (2011) suggested that the medial prefrontal cortex and the temporoparietal junction respond differently to a familiar person compared to a stranger. Moreover, the role of the anterior temporal lobe has recently been described as an important site for facial identity perception (Behrmann & Plaut, 2013; Collins & Olson, 2014). In the emotion perception domain, the amygdala and insula are involved in negative emotions like fear and disgust (Calder et al., 2007; van Ditzhuijzen, Keane, & Lawrence, 2007; Pessoa & Adolphs, 2010). Todorov and Engell (2008) also suggested that the amygdala, a subcortical structure that connects with the superior colliculus and pulvinar, provides continuous vigilance by evaluating faces and activates before interacting with them (also see Vuilleumier et al., 2001; Engell et al., 2007). Interestingly, although both the amygdala and the STS are heavily involved in the perception of facial expressions, they represent expressions differently (Harris, Young, & Andrews, 2012): the STS represents facial expressions in a continuous way, as its activation is insensitive to emotional change, while the amygdala represents in a categorically way, as its activation is 'sensitive to changes in expression that alter the perceived emotion'. Noticeably, some insightful research on the amygdala suggested that it might have a complicated role in the perception of facial expression. Patient SM, whose amygdala was lesioned, has been an important case for supporting the notion that the amygdala is vital in the perception of facial expression (Adolphs, 2006). However, a thorough study using the Bubble technique (Gosselin & Schyns, 2001) and eye tracking measurements has suggested that with proper instruction (e.g. 'look at the eye region'), patient SM could successfully distinguish fear as well as neural typical controls (Adolphs et al., 2005; Tranel, Schyns, & Damasio, 2005). Therefore, it is possible that the amygdala is the prerequisite (by guiding attention), rather than the processing hub for emotion perception.

Within the repertoire of face perception, this thesis has specifically focused on the perception of facial expressions and facial attractiveness for several reasons. Firstly, both facial expressions and facial attractiveness are vital for social communication. Facial expression reflects not only the current mood, but also an individual's possible intentions (Elias, Dyer, & Sweeny, 2017). On the other hand, facial attractiveness has a huge impact on mate choices as well as decisions about social exchange (Little, Jones, & DeBruine, 2011; Rhodes, 2006; Thornhill & Gangestad, 1999; Willis & Todorov, 2006). Studying ensemble coding of faces allows for a better understanding of face perception in a group context. Secondly, studying ensemble coding of faces could clarify whether ensemble statistics are prevalent in face perception. Are ensemble statistics ubiquitous in face perception?

Or is it just a phenomenon in the invariant aspect of face perception? So far, researchers are not fully clear about how ensemble statistics shape the perception of temporally presented faces. Moreover, there is no current evidence for the existence of ensemble statistics of facial attractiveness. Thirdly, studying the perception of facial expressions and facial attractiveness allows for a better understanding of the mechanisms of face perception and ensemble statistics. As facial expressions and attractiveness require different kinds of facial information (facial expression requires both local and holistic information, while facial attractiveness relies on holistic information) and employs distinctive circuits of neural systems, studying them enables us to: (a) clarify whether the ensemble statistics of faces relies on local or holistic information; (b) examine whether the ensemble statistics of faces relies on specific neural systems to calculate them, or if it relies upon distributed neural systems for face processing.

1.2.2. Facial Expression

Facial expression might be one of the most important elements in inter-person communication. Effective perception of facial expressions permits a direct recognition of others' current mood and also allows accurate speculation of their intention (Jack & Schyns, 2017).

Although there is an ongoing debate on what constitutes basic facial expressions (Ekman, 1993; Jack, Garrod, & Schyns, 2014; Jack & Schyns, 2017; Russell, 1994), there is more agreement among researchers as to the perception of facial expressions (Calder & Young, 2005). Similar to most facial characteristics, facial expressions require both feature and holistic processing. Holistic processing (Maurer, Le Grand, & Mondloch, 2002), sometimes termed global or configural processing, implies that the processing is not only on local features, but on the relationship among the features (the holistic information). Using a composite effect paradigm, Calder and colleagues

(2000) provided strong evidence for the 'holistic' claim. When two face-halves with different emotions were aligned, observers' performances were severely impaired when reporting the accurate emotion of either half compared to the misaligned conditions. This is because in the aligned condition the emotion perception of one half face is heavily affected by the other half, i.e. because of the holistic processing. However, Tanaka and colleagues (2012) found no evidence in support of the holistic processing of congruent expressions of happy and angry faces (the two face-halves with the same emotion). They concluded that the perception of facial expression utilizes both holistic and local information. Moreover, evidence from adaptation studies (e.g. Xu, Dayan, Lipkin, & Qian, 2008) further suggested that, at least for certain facial expressions, such as happy and sad, local features (e.g. mouth shape) determine perception of facial expressions.

Studies of the relationship between facial expression and other facial characteristics have produced mixed results. Several early studies in neuroimaging suggested that facial expression and facial identity are encoded in distinctive neural systems (e.g., Haxby et al., 2000; Kanwisher et al., 1997; Winston et al., 2004; see also Andrews & Ewbank, 2004). The 'independent' hypothesis was further supported by research in prosopagnosia: although prosopagnosics have difficulty in differentiating individual identity from faces, they have comparatively normal emotion perception (Humphreys, Avidan, & Behrmann, 2006 & Behrmann, 2006). However, recent evidence has suggested that the perception of facial expression and facial identity are *interdependent* (Calder & Young, 2005). For instance, Fox and Barton (2007) found adaptation to emotional faces generated significant aftereffects in test faces of another identity; however, the aftereffect was bigger with test faces with the same identity (not necessarily the same image) as the adapting face. However, the reverse is not true. In a follow-up experiment, they (Fox, Oruc, & Barton, 2008) suggested that identity adaptation is irrelevant to facial expression. Thus, there are both "identity-dependent" and "identity-independent" representations of facial expression. Interestingly, the perception of facial trustworthiness and attractiveness are also intertwined with that of facial expression. Using an adaptation paradigm, Engell and colleagues (2007) showed there is a common neural mechanism for the perception of facial trustworthiness and facial expression. Adaptation to angry expression makes subsequently viewed faces appear more trustworthy, while adaptation to a happy expression makes subsequently viewed faces appear less trustworthy. Moreover, the orbital frontal cortex (OFC) is activated by both attractive faces and face smile (O'Doherty et al., 2003 Perrett, Burt, & Dolan, 2003). Therefore, the perception of facial expression is substantially linked to other aspects of face perception.

1.2.3. Facial Attractiveness

Due to its social and evolutionary significance, facial attractiveness is one of the most important and widely studied facial traits (Little et al., 2011; Rhodes, 2006; Thornhill & Gangestad, 1999; Willis & Todorov, 2006). Research has uncovered numerous attributes that modulate facial attractiveness (Little et al., 2011; Rhodes, 2006), including: symmetry (Grammer & Thornhill, 1994; Perrett et al., 1999; Lee, Rowland, & Edwards, 1999; Rhodes et al., 2001), masculinity/femininity (sexual dimorphism; Perrett, Lee, Penton-Voak, & Rowland, 1998; Perrett, May, & Yoshikawa, 1994) and averageness (DeBruine et al., 2007; Deffenbacher, Vetter, Johanson, & O'Toole, 1998; Grammer & Thornhill, 1994; O'Toole, Price, Vetter, Bartlett, & Blanz, 1999; Perrett, May, & Yoshikawa, 1994; Rhodes & Tremewan, 1996; Rhodes et al., 2001; Valentine, Darling, & Donnelly, 2004). Moreover, such preference is universal (Cunningham, Roberts, Barbee, Druen, & al, 1995; Perrett et al., 1998; Rhodes et al., 2001). For instance, Perrett and colleagues (1994) showed that Japanese and UK subjects agree in regard to female attractiveness.

The facial attribute that appears to most determine facial attractiveness is the averageness. At the dawn of modern psychology, Galton (1878) reported that averaging leads to '…in every instance, a decided improvement of beauty' (Valentine et al., 2004). Contrary to the meaning of 'average', a face with high averageness is actually highly typical in the group with few idiosyncratic features (Little et al., 2011). Studies using computer-generated faces and real faces both suggest that individuals are more likely to prefer average faces to atypical ones (DeBruine et al., 2007; Deffenbacher et al., 1998; Grammer & Thornhill, 1994; Langlois & Roggman, 1990; Light, Kayrastuart, & Hollander, 1979; Perrett et al., 1994; Rhodes & Tremewan, 1996; Rhodes et al., 2001).

The preference for averaged faces is partly due to the fact that the average face is close to the face norm (center of the face space), and therefore preferred in visual processing (Little et al., 2011; Valentine, 1991; Valentine et al., 2004). However, the preference for averageness is also due to evolutionary reasons. Namely, genetic diversity, reflected by averageness, is itself appealing. Thornhill (1994) argued that the ‘owners’ of more average faces have more diverse sets of the genes and therefore more fitted for reproduction. Accumulated evidence also suggests that averageness is a good predictor of health (Rhodes, Maloney, Turner, & Ewing, 2007; Rhodes et al., 2001). Because of the importance of averageness, in an early work, Langlois (Langlois & Roggman, 1990) even (wrongly) claimed that “attractive faces are only average”. Apparently, facial attractiveness is not only about averageness (Langlois corrected her claim in her later papers). The earlier confusion is partly caused by the averaging algorithm used at that time: back then, controls in other facial attributes were always confounded with averageness. Using improved algorithms (Tiddeman, Burt, & Perrett, 2001), modern researchers have been able to differentiate the impact of symmetry and averageness on attractiveness (Burt & Perrett, 1997).

Research on averageness has revealed several interesting findings. Using a computer technique, Langlois and Roggman (1990) showed that average faces are more appealing than their components. Also, when more faces are averaged, the averaged face becomes more attractive: the average of 32 faces is much more attractive than the average of 2 faces. However, not all averaged faces are equally attractive (DeBruine et al., 2007; Perrett et al., 1994): the average of unattractive faces is still less attractive than that of attractive faces.

In addition to these intrinsic qualities within a face, external factors such as context and experience may shape the perception of facial attractiveness (Anderson, Lindner, & Lopes, 1973; Ewing, Rhodes, & Pellicano, 2010; Hönekopp, 2006; Jones, DeBruine, Little, Burriss, & Feinberg, 2007; Little, Burt, Penton-Voak, & Perrett, 2001; Little et al., 2011; Rhodes, 2006). For example, simply being viewed in the company of others helps us appear more attractive than when we are seen alone: the cheerleader effect (Walker & Vul, 2014). Recent exposure to an unattractive face also makes subsequently viewed faces more attractive: an attractiveness adaptation aftereffect (Rhodes, 2006).

1.2.4. Face Adaptation

Face perception is not always robust: recent experience can change our perception of subsequently viewed faces dramatically (Leopold et al., 2011; Webster et al., 2004; Webster, 2011). For instance, prolonged exposure to an emotional face can predispose the perceived emotion of successive faces: this is the face adaptation aftereffect (Webster, Kaping, Mizokami, & Duhamel, 2004; Ying & Xu, 2017). Though face adaptation shares some fundamental features with simpler visual adaptation, it has been broadly judged as high-level. Unlike simpler visual adaptation, face adaptation does not solely rely on the early-stage perception system like the retina or primary visual cortex; it occurs later within the core system of face processing (Leopold, Rhodes, Muller, & Jeffery, 2005; Webster & MacLeod, 2011).

Mechanisms of adaptation have been proposed according to different perspectives, neurally, psychologically and computationally. Two possible theories of visual coding have been widely discussed. One is norm-based opponent coding. The theory posits that there exist anti-neurons that prefer two opponents of the same characteristic of particular visual stimuli (e.g., male vs. female, face gender). Evidence from expression, identity and gender adaptation supports this theory (Burton, Jeffery, Calder, & Rhodes, 2015; Susilo, McKone, & Edwards, 2010). The other theory of adaptation involves exemplar-based, multichannel coding, and posits that neuronal populations have multiple channels to the different aspects of the same stimulus (e.g., the direction of eye gaze). Evidence from eye gaze direction (Seyama & Nagayama, 2006), head orientation and face viewpoint (Fang & He, 2005) has supported this theory (Calder, Jenkins, Cassel, & Clifford, 2008; Lawson, Clifford, & Calder, 2011). Despite the differences between the two theories of adaptation, however, both suggest that the "consequence of adaptation is normalization of the responses across mechanisms (Webster, 2015)". This normalization of the neural response will "calibrate for the mean stimulus level (Webster, 2014)". Adaptation processing stabilizes the visual system in changing surroundings by calibrating the coding mechanisms to maintain the optimized discriminative sensitivities (Yang, Shen, Chen, & Fang, 2011). We can thus predict that prolonged exposure to an extreme face, for example, a single happy face, will shift the face norm towards the adapting extreme face.

Several lines of research have shown that facial adaptation aftereffects can be produced in different features from low-level to high-level, including: curvatures (Xu et al., 2008), sexual dimorphism (Little, DeBruine, & Jones, 2005), eye-spacing (Little et al., 2005), identity (Leopold, O'Toole, Vetter, & Blanz, 2001; Rhodes & Jeffery, 2006), emotional expression (Hsu & Young, 2007; Luo, Wang, Schyns, Kingdom, & Xu, 2015; Webster et al., 2004; Xu et al., 2008), gender (Webster et al., 2004) and age (Schweinberger et al., 2010).

The adaptation paradigm has been widely employed in face research since it is a sensitive tool to "reveal the mechanism of face processing" (Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003). Using adaptation aftereffects, psychologists can probe the visual system "via the indirect but none the less revealing route of studying certain illusions (Frisby, 1979)". Because of the 'neat relationship which often exists' between the psychological findings using the adaptation paradigm and the neurophysiological findings using microelectrodes, some researchers called the adaptation paradigm 'the psychologist's microelectrode' (Frisby, 1979; Webster, 2015). For instance, researchers have studied the perception of adapting faces by measuring the magnitude of the adaptation aftereffects they have generated (Little et al., 2008; Luo et al., 2015; Rhodes & Jeffery, 2006; Webster et al., 2004).

1.2.5. Ensemble statistics

The visual system encounters a plethora of complex visual stimuli regularly. However, there is a huge discrepancy between the apparent rich environment and the limited processing ability of the human visual system (Cohen et al., 2016). The capacity and temporal limitations of working memory narrows perception onto a finite number of objects at any one time (Cowan, 2010; Haberman & Whitney, 2012; Nieuwenstein & Potter, 2006), resulting in the memory and recognition of an individual member of the a visual ensemble representation to be compromised. For instance, although grasping one component of an RSVP (Rapid Serial Visual Presentation) stream is possible (Keysers & Perrett, 2002; Potter, 2014), scrutinizing every stimulus in the RSVP stream is impossible.

The human visual system processes the stream of visual input with elegant strategies. Alvarez (2011) suggested that redundant and repeated features are omnipresent in visual experience: the real world is not random, but has structures, regularities, and similarities. Recent studies in ensemble statistics have high lighted a possible explanation of the processing of multiple stimuli: individuals might comprehend the surrounding world by obtaining the averaged gist using the ensemble statistics (Alvarez & Oliva, 2009; Haberman & Whitney, 2012).

Individuals derive the gist information of stimuli by averaging them. Such ensemble processing has been found in low-level features like location (Alvarez & Oliva, 2008), size (Chong & Treisman, 2005; de Fockert & Marchant, 2008; Li & Yeh, 2017), and high-level features like facial features (Haberman et al., 2009; Haberman & Whitney, 2010, 2012; Sweeny & Whitney, 2014). For example, Haberman and colleagues (2009) showed that after viewing a group of emotional faces, an individual could report the mean emotion of them with high accuracy. The studies (Haberman & Whitney, 2012) further suggested that such ensemble coding relied on holistic processing, not only on low-level feature processing: the accuracy of the reported mean emotion of the inverted faces was statistically smaller than that of upright ones. This averaging has many names, but is commonly known as ensemble coding (Alvarez, 2011; Ariely, 2001; Haberman et al., 2015; Haberman & Whitney, 2007, 2012; Whitney & Leib, 2017), ensemble statistics, or summary representations, and can occur both spatially (i.e., multiple faces presented at once in a scene; e.g., Haberman & Whitney, 2007) and temporally (i.e., different faces presented one at a time in rapid succession; e.g., Ying & Xu, 2017). Noticeable, most studies in ensemble face perception dealt with spatial ensembles, all faces presented at once. Therefore, temporal ensemble statistics require further investigation. Study 1 directly addressed this question.

Despite researchers widely describing ensemble statistics as extracting the gist of a scene, it is still far from clear what this gist represents (Whitney & Leib, 2017). For example, does ensemble coding extract a general representation, whereby a high-level judgment is made for each of the faces before summarizing the overall gist from these judgments? Or does the brain actually extract the low-level visual properties of each face first and then computationally average the facial deviations together? Here, Study 2 aimed to clarify this question.

Ensemble statistics are well observed in the changeable aspects of a given face. A recent study suggested that ensemble coding occurred during the perception of gaze viewpoints of a group (Sweeny & Whitney, 2014). However, the ensemble perception of invariant aspects of faces is only limited to facial identity. For example, Fisher and colleagues (2014) found that participants reported the mean identity of the crowd across multiple face viewpoints. Can we conclude that the ensemble statistic is a fundamental mechanism of visual perception? In this thesis, we examined the spatial and the temporal ensemble statistics of facial expressions and facial attractiveness to answer this question.

Counterintuitively, Leib and colleagues (2012) showed that developmental prosopagnosia (DP) cases were able to form ensemble representations of facial identity, despite their difficulty in identifying the individual identities. Therefore, authors suggested that the processing of a group of faces might differ from the processing of an individual face. However, this finding did not necessarily reflect the different processing mechanisms between sets of faces and static faces (Brady & Alvarez, 2011; Haberman & Whitney, 2009). Alternatively, as suggested by Alvarez

(2011), the ensemble statistics might smooth out the inner noise of the DP's representation of individual faces, and thus generate a comparable summarizing outcome to neural typical subjects.

1.2.6. Attention and Face Perception

The visual system is able to actively assign attention to a subset of sensory input, the processing of which will be facilitated by attention (Posner, 1980; Rensink, 2000; Treisman & Gelade, 1980). Attention, intertwined with all levels of visual processing, has been commonly believed to be of great importance in visual processing (Awh, Vogel, & Oh, 2006; Gazzaley & Nobre, 2012). The impact of attention has been observed behaviorally and neurally. Behaviorally, for instance, attention could lower contrast thresholds (Bisley & Goldberg, 2003) and decrease reaction times (Posner, 1980; Zhang, Wang, & Goldberg, 2014). Neurally, for instance, attention could enhance the firing rate of cells (Wurtz, Goldberg, & Robinson, 1982), as well as alter the BOLD signal (Vuilleumier et al., 2001). One effective way to modulate attention is distractor cueing (Bisley & Goldberg, 2003; Posner, 1980; Vuilleumier et al., 2001). It has been found that response times were shorter when subjects were cued to the spatial location of the stimuli compared to when cued away from the stimuli (Posner, 1980). Also, when guided by attentional cues to emotional faces rather than to houses, subjects exhibit significant emotion-related frontal positivity (Holmes, Vuilleumier, & Eimer, 2003). On the other hand, insufficient attention jeopardizes emotion perception (Adolphs, 2006; Adolphs et al., 2005; Holmes et al., 2003; Raymond, Fenske, & Tavassoli, 2003).

Visual attention is not only overt, e.g. linked to changes in gaze. Covert attention is vital in vision. Following the classic Posner cueing task (1980), researchers modified the cueing task, thereby further indicating that the reaction time (RT) differences between incongruent cue and congruent cue conditions reflected the strength of the ability to capture attention. From the classic Posner cueing task paradigm (1980) to many face adaptation studies (Luo et al., 2015; Rhodes & Jeffery, 2006; Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003; Xu et al., 2008) participants maintained their overt attention at the central fixation cross, while being able to sufficiently cast their covert attention onto periphery visual stimuli and to do the category tasks.

There is an interesting relationship between face perception and attention. Finkbeiner and Palermo (2009) found that facial stimuli could always capture attention, denoted by faster reaction times; however, other stimuli could not. On the other hand, Rhodes and colleagues (2011) showed that under different attention allocation conditions, individuals perceived faces differently: enhanced attention manipulated by tasks amplified the magnitudes of visual adaptation aftereffects. Therefore, utilizing attention allows the visual system to perceive faces in a better manner (Palermo & Rhodes, 2007; Vuilleumier et al., 2001). On the other hand, evidence from neural imaging experiments suggests that face processing in the amygdala is hardly affected by attention modulation, while cortical face processing is gated by spatial attention. For example, Vuilleumier and colleagues (2001) showed that the activation of FFA to facial stimuli was modulated by attention; however, the left amygdala responded to fearful expressions regardless of attention. Some evidence even suggests that the amygdala responds to various kinds of fear-related stimuli (e.g. fearful sound effect; Scott et al., 1997; Johnson, 2005); suggesting that the amygdala is a 'vigilant monitor'. Another experiment also showed that the emotion-related ERP differences from cortical regions were undetectable when the faces were outside of the attended location (Holmes et al., 2003).

Researchers have extensively studied ensemble representation as well as attention,. Some previous studies suggested that the ensemble coding of faces could be operated implicitly and explicitly (Haberman & Whitney, 2009; Ying & Xu, 2017) and could occur with reduced attention (Alvarez & Oliva, 2009). Recently, Elias and colleagues (2017) showed that attention is a prerequisite for perceptual averaging.

On the other hand, studies in ensemble representation of size have revealed that the impact of attention modulation on low-level ensemble coding (Chong & Treisman, 2005; de Fockert & Marchant, 2008; Li & Yeh, 2017). For instance, De Fockert and Marchant (2008) found that attended items contributed more to the ensemble representation. Therefore, the ensemble coding of size might occur in brain circuits which are subject to attention, perhaps the early visual cortex (Fang, Boyaci, Kersten, & Murray, 2008; Moran & Desimone, 1985). Thus, it is reasonable for one to postulate that the ensemble coding of faces might also be subject to attention modulation.

Chapter II: Study 1: Adaptation Reveals the Temporal Ensemble Coding of Sequentially Presented Emotional Faces

In real life, faces are usually temporally adjacent to other faces. How can we perceive faces presented in rapid succession? How could previous exposure to a stream of faces impact our face perception? Recent work on face perception has suggested that the visual system could interpret temporally presented faces via ensemble statistics. However, those findings were mainly based on the studies examining spatial ensemble statistics, the detailed perception of temporally presented faces is unclear. This study, with four experiments, systematically investigated the emotion perception of faces presented in RSVP streams using an adaptation paradigm.

2.1. Experiment 1.1: RSVP of emotional faces can generate adaptation aftereffects

Faces presented in a rapidly changing sequence can either be segregated for target detection or grouped for holistic representation. This experiment investigates whether passively viewing consecutively presented face sequences containing emotional information will generate facial expression aftereffects. We exposed subjects to a sequence of faces and required them to judge subsequently presented faces' emotion. The adapting faces were with the same emotion, but from different identities (either happy or sad). If the subjects did group the face streams into a holistic representation during adaptation, it is expected to observe a facial expression adaptation aftereffect comparable to other experiments.

2.1.1. Methods

Subject. Ten subjects (5 females, total mean age 22.8) with normal or corrected-to-normal vision participated in this experiment. Although one subject (HY) was the experimenter, the subjects were naïve to the purpose of the experiment. We collected written consent before the experiment. This study and the following ones were approved by the Internal Review Board (IRB) at Nanyang Technological University, Singapore, in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving human subjects.

Apparatus. The visual stimuli were presented on a 17-inch Philips CRT monitor (refresh rate 85 Hz, spatial resolution 1024 × 768 pixels). The experiment was controlled by Matlab R2010a software (Mathworks, MA) with Psychophysics Toolbox (Brainard, 1997; Pelli, 1997), on an iMac computer (Apple Inc., CA). Subjects rested their heads on the chin-rest located at a distance of 75 cm in front of the monitor in a dimly lit room. Each pixel subtends 0.024° on the screen.

Stimuli. The face stimuli were chosen from Karolinska Directed Emotional Faces (Lundqvist, Anders Flykt, & Arne Öhman, 1998) database. The KDEF database contains 4900 high quality pictures of human facial expressions. It is one of the most popular face databases (used in more than 2000 published papers) among face and emotion researchers. The models of the KDEF database are all Caucasian amateur actors. The pictures of ten identities with happy, sad and neutral expressions from the same identity were selected. We used the Webmorph software (DeBruine & Tiddeman, 2017) to manipulate and morph these faces.

Adapting Stimuli. The adapting stimuli were ten emotional faces from the KDEF face set. Happy and sad faces of the ten identities were chosen (Fig 2.1 A, B). The happy and sad faces were used because this pair of emotions has been widely used in facial expression adaptation studies (Fox & Barton, 2007; Hsu & Young, 2007), and has been believed to induce contrastive adaptation aftereffects (Xu et al., 2008). The adaptors were masked so that only the face region was visible during the experiment. All adapting stimuli were of a size of $2.40° \times 3.02°$.

Test Stimuli. We averaged all of the 35 neutral female faces of different identities from the KDEF face database to create a neutral face template. Similarly, we averaged 35 happy faces of these identities to generate the happy face template. We then morphed the neutral and happy templates to generate a sequence of test faces, varying in the proportion of happiness. The proportions of happiness in test faces are 0, 10, 20, 30, 35, 40, 45, 50, 60, 70% (Fig 2.1 C). Only the face region of test stimuli is displayed in the experiment. The test stimuli have the same size as the adaptors.

Figure 2.1. A. Adapting faces, ten happy faces from different identities used as adapting stimuli. B. Adapting faces, ten sad faces from different identities used as adapting stimuli. C. Test faces, ten morphed faces used as test stimuli.

Procedure

This experiment adapted subjects to the RSVP face sequences with different identities but the same emotion and measured the facial expression aftereffect. The adapting stimuli were either the happy or the sad RSVP face streams (examples in Figure 2.1A&B). In each trial, the test stimulus was one of the morphed faces (Figure 2.1C). The face stimuli and the central fixation cross were horizontally aligned (center-to-center distance of 3.5°), presented randomly to the right or left of the central cross.

There were three conditions in this experiment: happy RSVP adaptation, sad RSVP adaptation, and a baseline condition. We used a block design to run this experiment. In each block, each test face was repeated 10 times (random order). Each block lasted around ten minutes, and there was a ten-minute rest in between two consecutive sessions, to avoid carryover effects to the next block. For each subject, the order of the conditions was randomized. Each subject went through the whole experiment within the same day. Data collection started after sufficient practice trials (10–20).

Subjects pressed the space bar to initiate each block (Fig 2.2). They were asked to fixate at the fixational cross throughout the experiment. Each trial initiated with a 1494 ms (127 frames) fixation period. After that, the adapting RSVP face sequence appeared for 3764 ms (320 frames). In the adapting streams, each image lasted 23.5 ms (2 frames, temporal frequency at 42.5 Hz) on the screen and then was replaced by another picture at the same location with no interval. The 10 happy or sad faces each repeated 16 times in random order in the RSVP sequence. After a 506 ms (43 frames) interstimulus interval, a testing stimulus appeared for 400 ms (34 frames), masked by two 47 ms (4 frames) random Gaussian noise masks. This short duration was chosen to maximize aftereffects (Wolfe, 1984). The mask was displayed to reduce the effect of any afterimage. In the baseline blocks, there was no adapting face. In each trial, only one test stimulus was presented for 400 ms (34 frames). After that, subjects were reminded to report the emotion of the test face (pressing “H” for happy, “N” for not happy) by a 50 ms beep noise. After the response, the next trial began. No feedback was given.

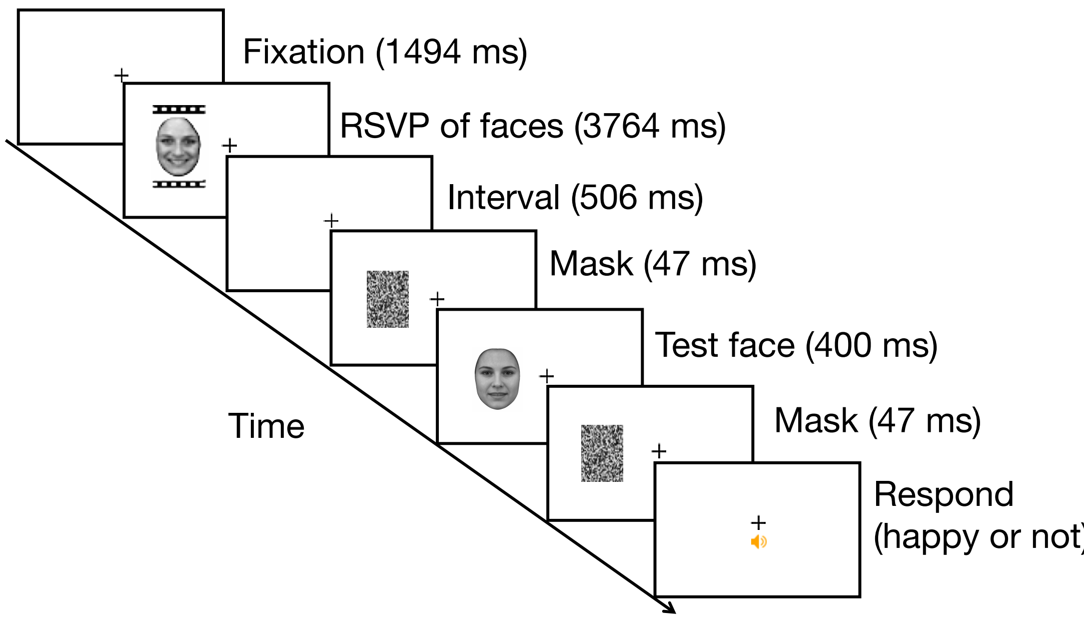

Figure 2.2. Trial Sequence of the happy RSVP adaptation condition. Subjects pressed the space bar to initiate one experimental block. After 1494 ms fixation, the RSVP stream of faces appeared for 3764 ms. After a 506 ms interstimulus interval (ISI), one testing face was presented for 400 ms, masked by two 47 ms random noise masks. The location of the test faces was identical to that of the RSVP stream of faces. The subjects was reminded by a 50 ms beep noise to report the test face’s expression (pressing “H” for happy, “N” for not happy). Experimental details are in the Methods section.

Data analysis

We plotted subjects' responses for each condition separately. The x-axis represents the emotion values of the testing faces; and the y-axis represents the proportion of the 'happy' response. We then fit these responses with this sigmoidal function $f(x) = 1/[1 + e^{-a(x-b)}]$. Here, "a/4" is the slope of the psychometic curve. The "b" is the parameter corresponding to the point of subjective equality (PSE: point of subjective equality; the 50% point of the psychometric function). Adaptation aftereffects were calculated by the PSE shifts from baseline condition. All statistical analyses were conducted in JASP 0.8.6 (JASP team, 2018), Matlab R2017a (Mathworks, MA, USA) and SPSS Statistics 22 (IBM, NY, USA).

2.1.2. Results

The results from a naïve subject are illustrated in Figure 2.3A. The proportion of the "happy" responses was plotted as a function of the proportion of happiness of the testing stimuli. The psychometric curve (solid black line) represents the baseline condition without adaptation. After prolonged exposure to the happy RSVP face stream, the psychometric curve (blue, Hap RSVP) shifted to the right: the subject was less likely the perceive the testing faces as happy. Similarly, in the Sad RSVP condition, the psychometric curve (red, Sad RSVP) shifted to the left: the subject perceived happy expressions more frequently. These psychometric curves closely resembled other experiments in facial expression adaptation (Hsu & Young, 2007; Webster et al., 2004, 2011; Webster & MacLeod, 2011). The new finding here is that during the prolonged exposure to the RSVP face stream, which contains multiple faces from different identities but the same expression (happy or sad), the subjects are able to extract the emotion from face stream involuntarily.

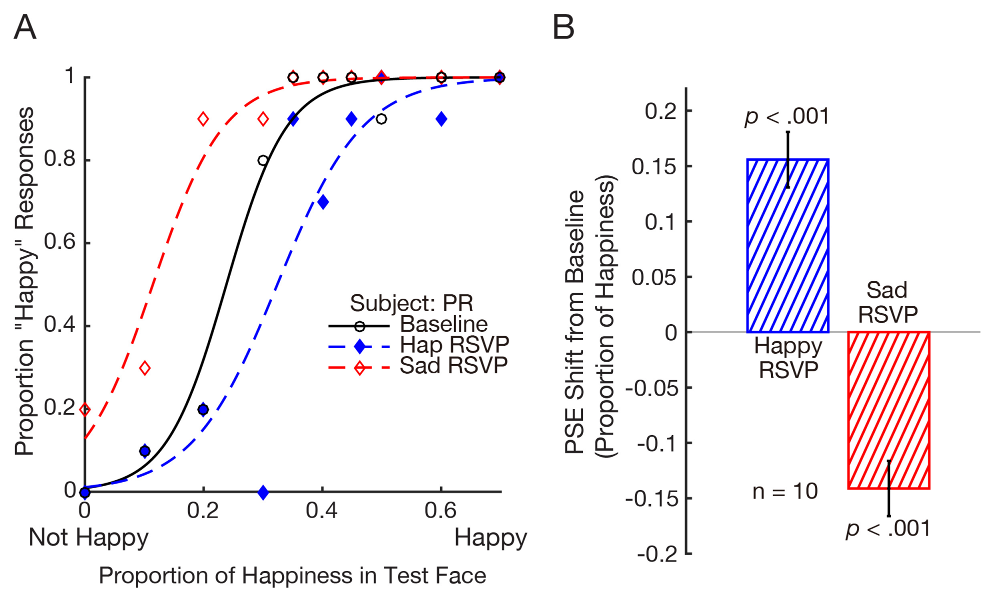

Figure 2.3. The effect of RSVP face stream adaptation (experiment 1.1). A, results from a naive subject PR. Baseline, the No adaptation baseline (solid black line). Hap RSVP, the happy RSVP face stream condition (dashed blue line). Sad RSVP, the sad RSVP Face stream condition (dashed red line). For each condition, we plotted the happy responses as a function of the fraction of "happiness" in the test face. B, Summary of data from all subjects. The averaged PSE shifts from baseline condition were illustrated (error bars indicate SEMs). The *p*-value shown above each bar in the figure were calculated by the paired sample *t*-test(two-tailed) between the adaptation condition and the baseline condition.

To quantify the facial expression aftereffects, we measured the PSE (the proportion of happiness corresponding to 50% happy responses) shifts away from baseline of all of the subjects (Figure 3B). The Shapiro-Wilk tests indicate that all data follow normal distributions, *p*s > 0.56. A positive PSE shift represents the psychometric curve (of the condition) is to the right of the "baseline" psychometric curve (less "happy" response). A negative value represents the psychometric curve (of the condition) is to the left of the "baseline" curve (less "happy" response). Both conditions generated significant aftereffects: significant positive aftereffect for happy RSVP adaptation condition (M = 15.579%, SEM = 0.0251; $t(9) = 6.219$, $p < .01$), and significant negative aftereffect for sad RSVP adaptation condition (M = -14.118%, SEM = 0.0249; $t(9) = -5.669$, $p < .01$).

2.1.3. Brief Discussion

The results here suggested that prolonged exposure to rapidly changing sequences (RSVP) of faces with the same emotion (happy or sad) generated significant facial expression aftereffects. This is interesting as in each RSVP stream during adaptation there were 160 faces from 10 different identities with the same emotion, with each face presented for only 23.5ms. Although the subjects were not required to do anything during the adaptation period, they were still able to extract the happy or sad emotional information in the RSVP stream during passive adaptation. This adaptation to the emotion in the RSVP stream then subsequently biased their judgment of the test faces, producing a facial expression aftereffect. It thus raises the question of how did this extraction of emotional information occur during RSVP adaptation?

Haberman and colleagues (2009) reported that participants perceived the sequential presentation of faces by averaging the faces in streams by ensemble statistics. Thus, the ensemble statistics may happen during adaptation to the RSVP stream in the current study. This begs the question as to whether adaptation to the faces in the RSVP stream would generate an equivalent facial expression aftereffect to an average face created from all of the RSVP faces.

2.2. Experiment 1.2: Ensemble statistics are involuntarily utilized during the RSVP of emotional faces

To test whether ensemble statistics are utilized during adaptation to an RSVP face stream, we adapted new subjects to the face average of the RSVP face stream and examined whether it generated similar facial expression aftereffects as adapting to the RSVP face stream. If they generate similar facial expression aftereffects, it would suggest that ensemble statistical averaging may occur during the adaptation stage of the RSVP face stream. If not, then this might suggest that other processes occur during adaptation, which remains yet to be further explored.

2.2.1. Methods

Subjects. Ten subjects (6 females, total mean age 22.1 years), with normal or corrected-to-normal vision, participated in experiment 2. Apart from the subject (HY) who was the experimenter, the other nine new subjects were naïve to the purpose of the experiment and different from the subjects in Experiment 1.1. All subjects gave written consent before testing.

Stimuli. All the visual stimuli were the same as in Experiment 1.1, except for two new face stimuli for adaptation. Using the Webmorph (DeBruine & Tiddeman, 2017), we created the average happy face from the 10 happy faces from different identities in the RSVP face stream; and similarly, we created the average sad face from the 10 sad faces from different identities in the RSVP face stream (Figure 2.4). These faces, with the same size as the testing stimuli, were used as the adaptors in the Experiment.

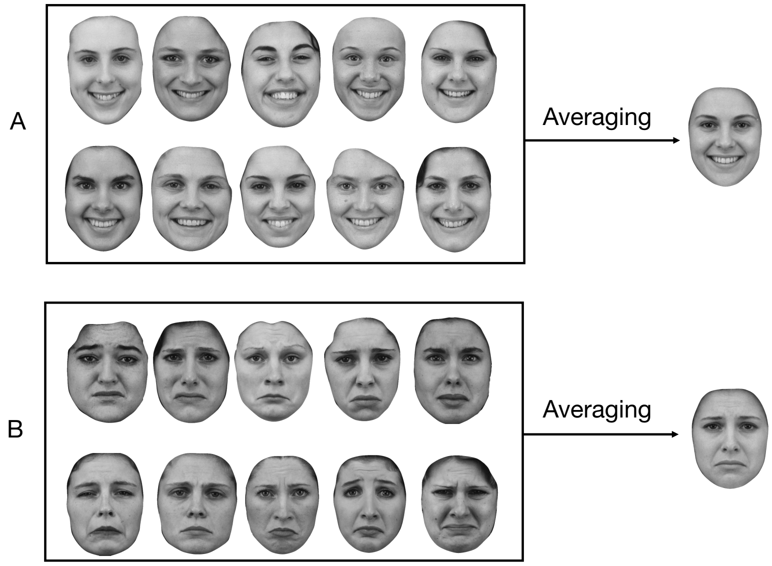

Figure 2.4. The averaged faces used as adaptors. A. The averaged face based on all ten happy faces. B. The averaged face based on all ten sad faces.

Apparatus, Procedure, and Data Analysis

The apparatus and data analyses were adapted from Experiment 1.1. The general procedure and the trial sequence were similar to Experiment 1.1, except for two new conditions. In this experiment, we added two additional adaption conditions: happy static adaption and sad static adaption. The happy static adaptation used the above generated happy averaged face as the adaptor (Fig 2.4A), and the sad static adaptation used the sad averaged face as the adaptor (Fig 2.4B). The static face adaptors were displayed for 3.82s on the screen in each trial, with the same duration as the entire RSVP sequence. Therefore, there were five conditions in total: happy static adaptation, sad static adaptation, happy RSVP adaptation, sad RVSP adaptation, and baseline (no adaptation) conditions. The subjects went through all five conditions in randomized blocks. Within each block, the test faces were randomly selected from the test face set (Figure 1c) for each participant.

Each test face was repeated 10 times in the experiment; however, each test stimulus was presented 20 times in Experiment 1.2 (in two different sections). Subjects finished the experiments on two different days within three consecutive days. On the first day, the subjects were tested on the 5 conditions in randomized blocks with 10 repetitions of each test stimuli; and the same occurs on the second day of the test. The testing orders were counterbalanced.

2.2.2. Results

We illustrated the results from a naive subject XY in Figure 2.5A. After exposure to the averaged happy face, the subjects were less likely to perceived happiness in the test faces. The psychometric curve (solid blue, Hap Static) shifted to the right, as it was for adapting to the happy RSVP face stream (dashed blue, Hap RSVP). Similarly, after adapting to the sad averaged face, the subjects were more likely to perceive happiness in the test faces. The psychometric curve (solid red, Sad Static) shifted to the left, as the subjects were adapting to the sad RSVP face stream (dashed red, Sad RSVP). This is interesting as it provides the first direct evidence that an average face generates similar adaptation aftereffects as its RSVP face stream.

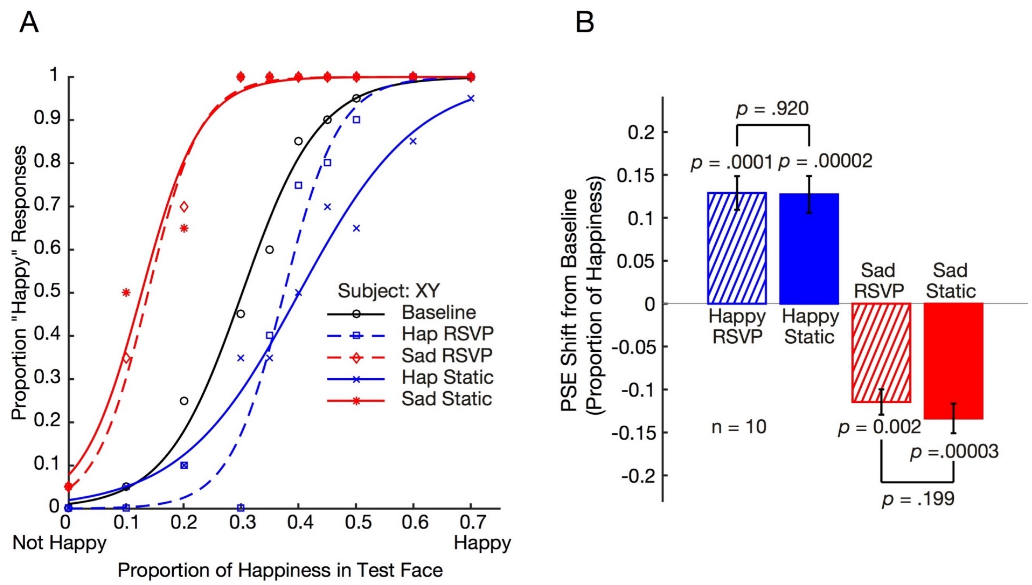

Figure 2.5. The effect of RSVP face stream adaptation and its statistically average face adaptation (experiment 1.2). A, a naive subject's psychometric functions under the following conditions: Baseline, No adaptation baseline (solid black line); Hap RSVP, adaptation to the happy RSVP face stream (dashed blue line); Sad RSVP, adaptation to the sad RSVP Face stream (dashed red line); Hap Static, adaptation to the static average happy face (solid blue line); Sad Static, adaptation to the static average sad face (solid red line). For each condition, the "happy" response was plotted as a function of the proportion of "happiness" of the test face. B, Summary of data from all subjects. The averaged PSE shifts from baseline condition were illustrated (error bars indicate SEMs). The *p*-value shown above each bar in the figure were calculated by the paired sample t-test(two-tailed) between the adaptation condition and the baseline condition.

The summary of PSE shifts from all ten subjects is shown in Figure 2.5B. The Shapiro-Wilk tests indicate that all data follow normal distributions, *p*s > 0.27. Compared to the baseline, all the adaptation conditions generated significant facial expression aftereffects: happy RSVP adaptation condition (M = 12.894%, SEM = .0196; $t(9) = 6.586$, $p < .01$); sad RSVP adaptation condition (M = -11.455%, SEM = .0147; $t(9) = -7.744$, $p < .01$); happy static adaptation condition (M = 12.722%, SEM = .0215; $t(9) = 5.929$, $p < .01$); and sad static adaptation condition (M = -13.365%, SEM = .0173; $t(9) = -7.743$, $p < .01$). Further comparisons of the RSVP face stream and average static face adaptation conditions revealed that there is no statistical difference between them: happy RSVP vs. happy static adaptation conditions ($t(9) = -0.103$, $p = .920$); and sad RSVP vs. sad static adaptation conditions ($t(9) = -1.388$, $p = .199$). Moreover, the RSVP adaptation generated similar magnitudes of aftereffects in Experiment 1 and 2: happy ($t(18) = 0.844$, $p = .409$) and sad ($t(18) = -0.920$, $p = .370$) RSVP adaptation. This thus suggests the findings from the experiments are robust and replicable.

Finally, we studied the relationships between the aftereffects caused by the RSVP face stream and its averaged face (Fig 2.6). We found significant correlations between the happy RSVP face stream and its average happy static face adaptation (*r* = .67, *p* = .033), and between the sad RSVP face stream and its average sad static face adaptation (*r* = .64, *p* = .046). The similar trend of correlation between the two emotions suggests the existence of the same mechanism in the static and RSVP face stream adaptation.

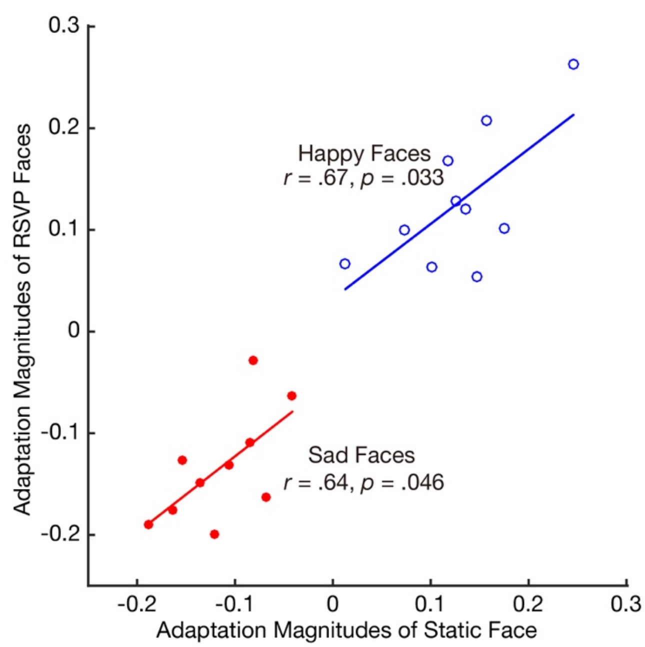

Figure 2.6. The magnitudes of adaptation evoked by static faces as a function of the magnitudes of adaptation evoked by RSVP faces. Ten blue dots represent (upper right corner) ten subjects' performances on two happy adaptation conditions, and ten red dots (lower left corner) represent ten subjects' performances on two sad adaptation conditions. For each emotion, the subjects' adaptation magnitudes of static and RSVP conditions were significantly correlated.

2.2.3. Brief Discussion

We found that adaption to the ensemble statistical average of the RSVP face stream generated similar facial expression aftereffects as adaption to the RSVP face stream. The magnitudes of the two aftereffects were significantly correlated. This finding supported our hypothesis and indicated that the subjects were automatically averaging the face stream in ensemble statistics during adaptation, even in a passively viewing condition. Studies on spatial visual integration, presenting a set of faces simultaneously, have demonstrated that the visual system interprets this set of faces by automatically averaging the spatially distributed faces together (Haberman et al., 2015). Our current experiment and previous studies support the view that the subjects perceived the average emotion of the face streams holistically in an ensemble statistical way (Haberman & Whitney, 2012). However, we provide the first direct evidence in adaptation that if we are exposed to a heterogeneous sequence of information (different identities), we are able to integrate the information and extract the common information (e.g., happy emotion), and the information will subsequently bias our judgement in the relevant tasks (e.g., emotion judgment). This process during adaptation is automatic and implicit, as we did not instruct the subjects to integrate and extract such information during adaptation.

However, all the face images in the RSVP stream are of the same emotion. Consequently, the average face and its individual faces are of the same emotion as well. It thus raises this following question: does adaptation rely on the emotion of individual faces or the average face of the RSVP stream? To answer this question, following experiments manipulated the emotions in the face sequence. Experiment 1.3 introduced variability to the adapting face sequence emotion, and Experiment 1.4 introduced a face with the opposite emotion into the RSVP face sequence.

2.3. Experiment 1.3: The temporal ensemble statistics are

determined by the mean, but not the variance

To manipulate the variability of the emotions of the adapting face sequence, we generated the RSVPs from faces showing various degrees of happiness, from 0.6 to 1.0 in the proportion of happiness, but maintaining the average RSVPs emotion as 0.8. There are three types of emotion variance during adaptation: varying from 0.6 to 1.0 randomly; 0.6 or 1.0 binomially; and all faces with 0.8 proportion of happiness. In addition, we also varied the temporal frequency of the RSVP stream, with high (42.5 Hz) and low (5.3 Hz) temporal frequency. Therefore, there are 6 additional adaptation conditions in this experiment. All of them have the same average face from the RSVP sequence but with different variations in emotion or temporal frequencies.

If ensemble coding occurs during adaptation, we would expect the same facial expression aftereffect (FEA) for all conditions, as they have the same average face during adaptation. However, if variation instead of ensemble coding counts in the RSVP adaptation, we would observe different FEA due to the variability among different conditions.

2.3.1. Methods

Subjects. Ten subjects (3 females, total mean age 22.4 years), with normal or corrected-to-normal vision, participated in the experiment. One of the subjects (HY) was the experimenter; the other subjects were naïve to the purpose of the experiment and different from the subjects in the previous two experiments. All subjects gave written consent before testing.

Stimuli. All the test stimuli and the raw adapting stimuli were the same as in Experiment 1 & 2. For the additional adapting faces, we morphed between the happy and neutral faces of each of the ten identities and generated 10 sets of faces, via the Psychomorph software online (DeBruine & Tiddeman, 2017). In each face set, there are 11 faces, with various proportions of happiness, specifically at 60, 64, 68, 72, 76, 80, 84, 88, 92, 98 or 100% (Fig 2.7A). Therefore, there are 110 faces (10 sets x 11 faces) in total. These faces were randomly selected in the RSVP face sequence based on the experiment condition. In addition, we morphed all of these 110 faces into one average face with 80% happiness. We then used this 80% happiness face as the adaptor in the static adaptation condition.

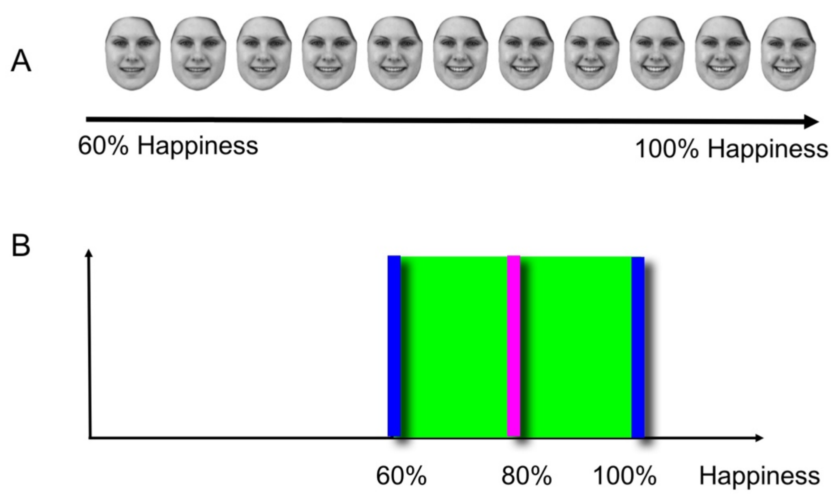

Figure 2.7. The adaptors. **A,** Examples of emotional faces used as adaptors. These faces were created by averaging the neutral and the happy faces of one facial identity. **B,** Schematic demonstration of the adapting faces selected on a hypothetical face space. The magenta bar indicates the faces chosen for the 80% condition, the blue bars indicate the 60or100% condition, and the green square indicates the 60to100% condition.

Procedure, and data analysis.

The general trial sequence and data analysis were adapted from the previous experiment, except for the new conditions: 6 adaptation conditions, static average face adaptation, and baseline. These 6 RSVP adaptation conditions vary in two dimensions: (1) emotion variance, and (2) temporal frequency. There were two temporal frequencies: high temporal frequency, with each face displayed on the screen for 23.5ms (42.5Hz); and low temporal frequency, with each face displayed for 188ms (5.3Hz). There were three types of emotion variances (see Fig 2.7B): (1) 60to100% condition, in which the adapting RSVP stream contains all of the adapting faces except for the 80% face, presented in random order; (2) 60or100% condition, in which the adapting face stream contains 20 individual faces showing either 60% or 100% happiness; and (3) 80% condition, in which the adapting face stream contains 10 individual faces each showing 80% happiness.

In each condition, each test face was randomly presented 20 times. To minimize the fatigue effect, we halved every condition into two identical blocks (similar to Experiment 2). Within each block, the orders of conditions were randomly selected.

2.3.2. Results

We plotted a naïve subject HN's results in Figure 2.8a. After adapting to all conditions (the 7 colored lines), the subject was less likely to judge the test face as happy, which was quantified by the direction of the psychometric curves shift away from that of the "baseline" condition (solid black line).

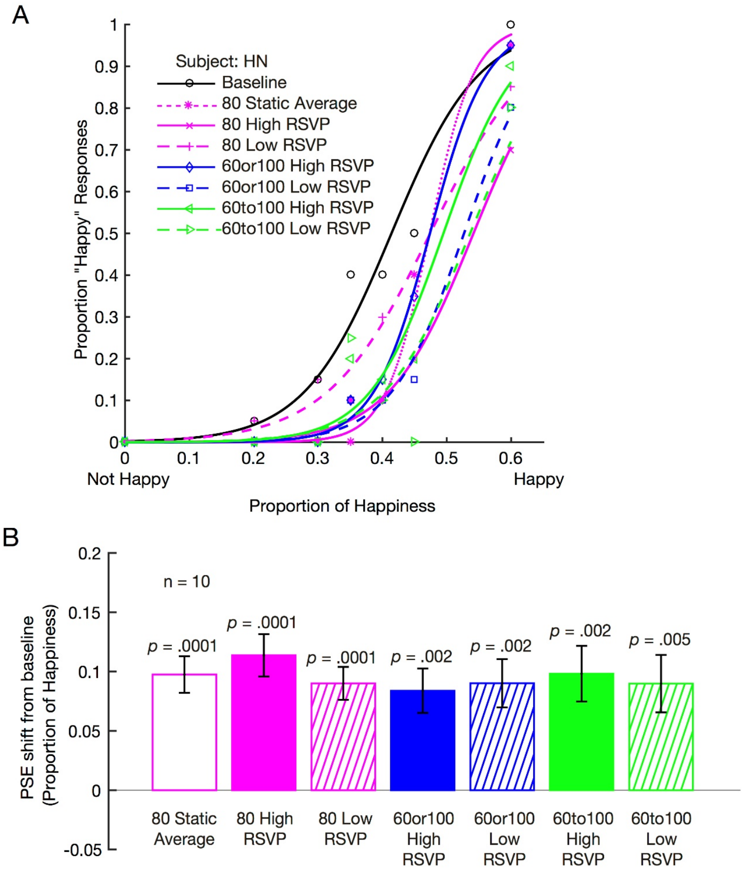

Figure 2.8. The effects of various adaptation conditions. **A**. The results from one naïve subject HN. Baseline represents baseline condition without adaptation (black line). 80 static averaged represents the static adaptation condition with the averaged faces (dashed magenta line). The other six lines represent the RSVP conditions. The full lines represent the high temporal frequency condition, while the dashed lines represent the low temporal frequency condition. Magenta, blue, and green lines represent RSVP streams consisting of faces showing 80% happiness, either 60 or 100% happiness, and randomly from 60 to 100% happiness respectively. For example, the full magenta line (80 H) indicates the condition with faces streams at a high temporal frequency, comprising with 80% of happiness. For each condition, we plotted the happy response as a function of the proportion of happiness within each testing face. **B.** Data summary. The white bar (magenta line) indicates the 80 static averaged condition. The full-color bars indicate high temporal frequency conditions (42.5Hz), and the hatched bars indicate the low temporal frequency conditions (5.3Hz). Magenta, blue, and green bars represent RSVP

streams consisting of faces showing 80% happiness, either 60 or 100% happiness, and randomly from 60 to 100% happiness respectively. For example, the full magenta bar with caption 80 High RSVP represents the 80% condition at the high temporal frequency. The averaged PSE shifts from baseline condition was illustrated (error bars indicate SEMs). The p-values shown above each bar in the figure were calculated by the *t*-tests and Wilcoxon rank sum test between the adaptation conditions and the baseline condition.

The adaptation aftereffects were calculated by the PSE shifts between the adaptation condition and the baseline. The summary of PSE shifts from all 10 subjects is shown in Figure 8B. The Shapiro-Wilk tests indicate that all data, but the '60to100 High RSVP' condition (p = .006), follow normal distribution, ps > 0.20. To investigate the effects of emotion variance and temporal frequency on RSVP adaptation, we performed a two-way 2 (temporal frequency) by 3 (emotion variance) Bayesian Repeated Measures ANOVA (Wagenmakers et al., 2017) on the facial expression aftereffects (PSE shifts from the baseline) on all six RSVP conditions. The results supported the null hypothesis over the main effect of temporal frequency (BF_{01} = 3.00; suggesting the null hypothesis is 3 times more likely to be true), emotion variance (BF_{01} = 4.7), and the interaction between the two factors (BF_{01} =

13.99). These suggest that neither emotion variance nor temporal frequency of the RSVP sequence plays an important role in RSVP sequence adaptation. Then how do these 6 adaptation conditions compare to the static average face adaptation?

A Friedman test showed that there was no significant difference among any of 7 the adaptation conditions ($\chi^2(2) = 2.5$, $p = .87$). Moreover, all seven adaptation conditions showed significant adaptation aftereffects (all $M > 8.38\%$, all $t > 3.710$, all $p < .005$; for the ‘60to100 High RSVP’ condition, Wilcoxon rank sum: $Z= 55$, $p = .002$). Further comparisons between the static adaptation condition with each of the RSVP conditions revealed no significant difference (all $t < .889$, all $p > .397$; for the ‘60to100 High RSVP’ condition, Wilcoxon signed-rank $Z= 34$, $p = .56$). These results thus suggest that adapting to the RSVPs of faces with variable emotions generated equivalent aftereffect as adapting to the averaged face of the RSVP sequence.

2.3.3. Brief Discussion

We found that varying the temporal frequency (high vs. low) or varying the strength of emotion in the RSVP sequence with the same average face does not change the facial expression aftereffects (FEAs) accordingly. They all generated the same adaptation aftereffects as their face average. Therefore, regardless of the emotion variance and temporal frequency changes in the RSVP sequence, as long as the face sequence has the same average face, it will still generate the same magnitude of FEAs.

Unlike Experiment 1.2 in which all of the adapting faces conveyed emotions at their maximum intensity, here we presented faces with different proportions of happiness. In the *60or100%* and *60to100%* conditions, the adaptation aftereffects were not reduced by the presence of the faces with 60% happiness in the adapting face sequence; nor increased by the presence of 100% happiness faces in the face sequence. Instead, they both generated similar facial expression aftereffects as the mean emotion (80%) of the face sequence. Moreover, in these two conditions, the mean emotion was not directly accessible; therefore, the substantial adaptation aftereffects could only be attributed to the ensemble statistics. Taken together, this experiment suggests that the strength of the aftereffects of the RSVP stream adaptation was not determined by the variability of the face stream but were determined by the mean emotion of the stream.

2.4. Experiment 1.4: Emotion, but not identity, determines the adaptation aftereffects

All of the above experiments tested the emotion of average face identity. To independently manipulate the average emotion and identity in the stream, we matched the identity of the test face with one identity from the adapting RSVP face stream. If individual faces in the face stream play an important role in adaptation, since the identity of the test face matched with one particular identity face in the face stream, we would expect to see the adaptation aftereffect influenced by the particular identity. On the other hand, if the pooled average of the face stream plays a more important role, it thus confirms ensemble representation during adaptation.

To single out a particular identity's faces, we manipulated its emotion to be the opposite of the other faces in the face stream. For example, if the face stream contains 9 happy faces and 1 sad face, the test faces and the sad face have the same identity. The average of the face stream is happy. If the subjects' judgment of test faces is influenced by the emotion of the same identity (e.g., sad) in the adapting face stream, they will tend to judge the test faces as happier. If the subjects' perception of test faces is determined by the emotion of the averaged face (e.g., happy) during adaptation, they will make more sad judgments. Therefore, the direction of the facial expression aftereffect will tell us whether adaptation is based on the emotion of a single face with the matched identity or on the averaged emotion of the entire face stream.

2.4.1. Methods

Subjects

Ten subjects (2 males, mean age 21.9 years), with normal or corrected-to-normal vision, participated in this experiment. One of the subjects was the experimenter (HY). The other subjects were naïve to the purpose of the experiment and different from the subjects in previous experiments. All subjects gave written consent before testing.

Stimuli, Apparatus, Procedure, and Data Analysis

The happy and sad faces from Experiment 1.2 were used in this experiment. We generated three new RSVP streams for the new adapting stimuli: 90% happy, 50% happy, and 10% happy in the face stream. In the 90% happy condition, 9 happy faces and one sad face was selected (see Fig 2.9A). Each face was randomly presented 16 times in the stream (at 42.5 Hz). In the 50% happy condition, we randomly selected 5 happy faces and 5 sad faces to create the stream. In the 10% happy condition, one happy face and 9 sad faces were used. Notably, the identity F07 always showed the opposite emotion to the majority of the stream in 90% and 10% condition, and it appeared as happy and sad randomly in the 50% condition. To minimize the low-level adaptation, the adaptors were at 133% of the size of test stimuli (Burton et al., 2015; Zhao & Chubb, 2001).

The test stimuli were generated from one specific identity (F07), which has been used in the previous RSVP streams. We morphed between the happy and neutral emotional faces of this identity and generated ten sets of faces, by the Psychomorph software online (DeBruine & Tiddeman, 2017). The proportions of happiness of the test stimuli were at 0, 10, 20, 30, 40, 50, 60% (Fig 2.9B). The size of the test stimuli was $2.40° \times 3.02°$, the same as those in previous studies.

Figure 2.9. Stimuli used in Experiment 4. A, The adapting stimuli used for 90% happy condition. This stream contained 10 faces (each repeated 16 times in random orders): 9 happy faces, and 1 sad face (in the red square). The sad face was from the same identity (F07) as the test stimuli. B, Testing stimuli. These were all generated from one chosen identity (F07).

The apparatus, general trial structure, and data analysis were adapted from Experiment 1.3. There are 4 conditions in total, 3 RSVP adaptation conditions and baseline. Each test stimuli was tested 20 times randomly in two separate blocks. Therefore, the four conditions were tested in 8 blocks with a ten-minute rest in between.

2.4.2. Results

A naïve subject HD's results are illustrated in Figure 2.10A. The blue dashed line represents adaptation to the 90% happy face stream and shifts the psychometric curve to the right of the baseline condition (black solid line). This indicates more sad judgments after adapting to the 90% happy face streams, despite the fact that the testing face is of the same identity as the 10% sad face in the face stream. The red dotted line represents adaptation to the 10% happy face stream (with 90% sad faces in the face stream) and shifts the psychometric curve to the left of the baseline condition. This indicates more happy judgments after adapting to the 10% happy face stream. The magenta dash-dotted line represents adaption to the 50% happy face stream; the psychometric curve did not shift much from the baseline condition. It thus indicates little judgment bias in adapting to 50% happy and 50% sad face streams.

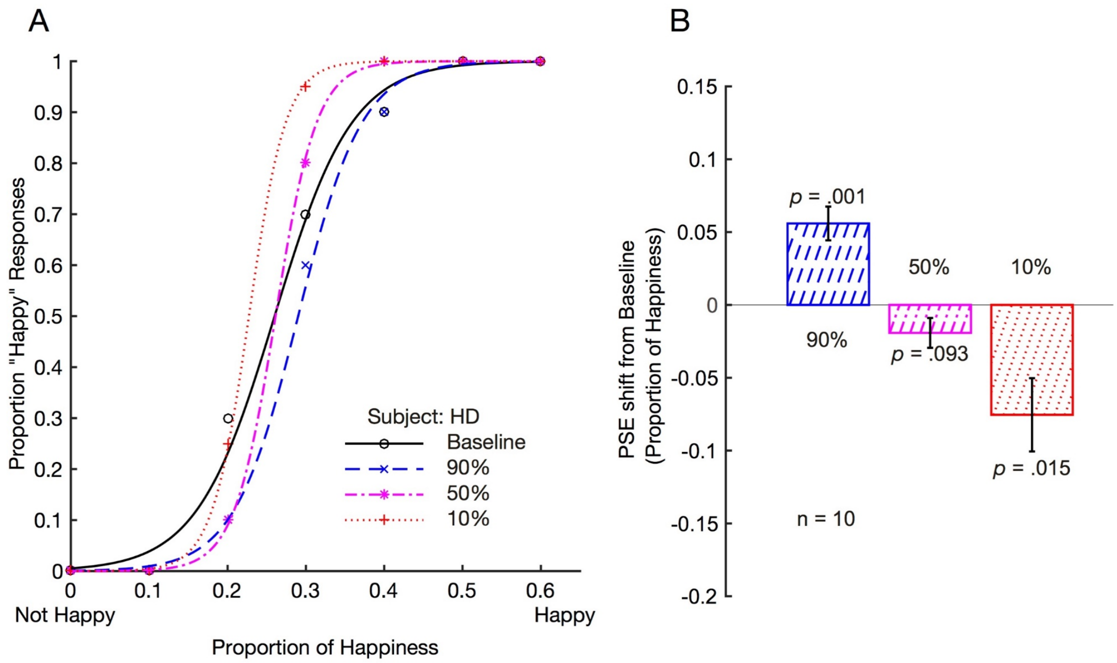

Figure 2.10. The effects of RSVP face streams with different proportions of happy faces. A, The results from a naïve subject HD. Baseline represents the baseline condition without adaptation (black line); 90% represents adapting to RSVP streams with 90% happy faces (blue line); 50% represents adapting to RSVP streams with 50% happy full faces (magenta line); 10% represents adapting to RSVP streams with 10% happy full faces (red line). For each condition, the happy response was plotted as a function of the proportion of happiness of the test face. B, Data summary from all ten subjects. The averaged PSE shift from baseline condition was illustrated (error bars indicate SEMs). The p-value shown above each bar in the figure was calculated by the paired sample t-test(two-tailed) between the adaptation condition and the baseline condition. Blue, magenta, and red bars represent the mean adaptation aftereffects under 90%, 50% and 10% of happy faces conditions correspondingly.

The summary of all 10 subjects' results is presented in Figure 2.10B. The Shapiro-Wilk tests indicate that all data follow normal distributions, *p*s > 0.30. A one way ANOVA found that different RSVP streams generated different adaptation aftereffects ($F(2,18) = 21.083$, $p < 0.001$, $\eta_p^2 = .701$), compared to baseline. Specifically, adapting to a 90% happy face stream generated significant aftereffects ($M = 0.0559$, $t(9) = 4.831$, $p = .001$); adapting to the 10% happy face stream also generated significant aftereffects ($M = -0.0754$, $t(9) = -3.003$, $p = .015$). The aftereffects of the two are in the opposite directions. Adapting to the RSVP streams with 50% happy faces and 50% sad faces did not generate significant aftereffects ($M = -1.92\%$, $t(9) = -1.878$, $p = .093$). This suggests that adaptation aftereffects are determined by the mean emotions of the RSVP streams of emotional faces, instead of the individual faces that are of the same identity as the test faces.

2.4.3. Brief Discussion

This experiment showed that adapting to a face stream with a majority (e.g., 90%) of happy faces still produces significant facial expression aftereffects in a similar way as adapting to happy faces. Matching the identity of the test faces to the minority (e.g., 10%) sad faces in the face stream did not affect the aftereffects in the same way as adapting to the sad faces. This is because the representation of the faces is defined by the overall emotion of the stream, but not the facial identity. Adapting to the 50% happy and 50% sad face stream did not produce any significant facial expression aftereffects, as the average emotion of the face stream is neutral. These results offer further evidence to support our hypothesis that ensemble coding occurs during adaptation to an RSVP face stream.

When we matched the identity of the test stimuli to one of the identities in the RSVP stream, but presented it with the opposite emotion to the majority of the other faces in that stream, the adaptation aftereffect complied with the majority emotion instead of that of the same identity. During the RSVP face stream, the subjects were able to extract the mean emotion of the face stream, instead of focusing on the emotion of a particular identity which was the same identity as the test faces. It thus suggests that a substantial ensemble representation of the gist emotion occurs when passively viewing a stream of emotional faces.

2.5. Discussion

This study investigated the effect of emotion adaptation to faces presented in RSVP streams. Experiment 1.1 showed that adapting to a set of faces with different identities but the same emotion (happy or sad) presented temporally affected judgments of emotion in subsequently presented faces. We then generated average faces from the adapting face sets. Experiment 1.2 demonstrated that the adaptation aftereffects generated by the RSVP face streams are equivalent to, and correlated with, the aftereffects generated by their average static face adaptation counterparts. Experiment 1.3 found that regardless of the variations in temporal frequency or emotion, the RSVP streams induced the same magnitude of aftereffect as the static averaged face of the streams. Experiment 1.4 further revealed that when presenting both happy and sad faces in the RSVP stream, the adaptation aftereffect is not dependent on the identity of the test faces, but dependent on the mean emotion in the stream: when the faces are half happy and half sad, there are no adaptation aftereffects; when 90% is happy, and 10% is sad, aftereffects are generated in the same direction as adapting to a happy face; and vice versa. These results together suggest that we are able to passively average the RSVP faces' emotions during adaptation. It further indicates that we can involuntary process the information into the ensemble representation even when we are not instructed to do so. The findings also delineated the possible mechanism for the formation of our facial expression norm.

The results in all experiments suggest that ensemble statistics are extracted automatically during adaptation. When the subjects were adapted to the RSVP of faces, the subsequent aftereffect was influenced by faces presented in the RSVP stream; the similar aftereffects from adapting to the RSVP face sequence and its static average face suggests that the subjects involuntarily integrate all the facial information together to create an ensemble representation. Moreover, strong correlations between the aftereffects of the RSVP of faces and the static faces' observed in Experiment 1.2, as well as the similar adaptation aftereffects between RSVP streams and the static averaged faces in Experiment 1.3, show that there are similar mechanisms between the processing of RSVP of faces and the static faces, i.e., the magnitude of RSVP face stream adaptation could be predicted from the magnitude of the static face adaptation (Haberman et al., 2015); thus further supporting the notion that the visual system interprets the RSVP of faces as the statistically averaged static face (Haberman et al., 2009).

All four experiments here suggest that adapting to a stream of faces of different identities generates a substantial and significant facial expression aftereffect, and this aftereffect is determined by the averaged emotion of the face stream. Previous studies have shown that after adapting to one identity posing a certain emotion, participants will then produce robust emotion adaptation aftereffects when judging emotion in subsequently presented faces of different identities; in other words, there is identity-independent emotion processing mechanism (Campbell & Burke, 2009; Fox & Barton, 2007). Noticeably, the test faces in the first three experiments were the average face from the adapting face stream. Therefore, the adapting and testing faces were from different identities. Moreover, the manipulation in Experiment 1.4 confirmed that it is the emotion, instead of the identity of the face stream, that affects facial expression aftereffects. This independence of facial identity and emotion is supported by models of face perception that posit the existence of distinct cortical regions involved in the separate processing of facial identity and emotion (Haxby & Gobbini, 2011; Haxby et al., 2000; Haxby, Hoffman, & Gobbini, 2002). Therefore, our current findings in adaptation aftereffects of RSVP are in line with these previous studies indicating identity-independent facial emotion processing.

Recent studies reported that four-year-old children are able to use certain ensemble processing (Sweeny, Wurnitsch, Gopnik, & Whitney, 2014), and could use norm-based coding for faces (Jeffery, Read, & Rhodes, 2013). Our study, together with these studies (see also, Rhodes, Neumann, Ewing, & Palermo, 2015), highlight a potential link between ensemble and norm-based face processing. This study delineated the formation of the facial expression norm during adaptation: our vision system implicitly integrates the multiple faces we encounter over time to the average face of the face stream, an ensemble representation updated norm of the facial expression space shaped by recent visual experience. Therefore, our findings, for the first time, offer new insight into the updating procedure of the facial expression norm and automatic face processing.

Chapter III: Study 2: Distinctive Mechanisms for Temporal and Spatial Ensemble Coding of Faces

When we are presented with an array of faces in a scene, our brains involuntarily extract the ensemble statistics of the information that they convey (Alvarez, 2011; Haberman & Whitney, 2007, 2012; Whitney & Leib, 2017). Despite researchers widely describing ensemble statistics as extracting the gist of a scene, it is still far from clear what this gist represents (Whitney & Leib, 2017). For example, does ensemble coding extract a general representation, whereby the high-level perception of faces (gist) is summarized? Or does the brain actually extract the low-level visual properties of each face first and then computationally average the facial deviations together?

Facial characteristics displayed in previous ensemble coding research have failed to produce a consensus as to which hypothesis is correct. For example, facial emotion predicts the same perceptual outcome for two competing hypotheses (Figure 1.1A). Therefore, using emotional facial expressions is not ideal to elucidate the possible computational mechanisms of ensemble coding. Here, we take advantage of the fact that a computational average face created from the low-level properties of a group of faces is more attractive than the gist of the underlying faces in the group (DeBruine et al., 2007; Perrett et al., 1994). By requiring participants to perceive facial attractiveness in a temporal ensemble fashion, we can clearly test for the first time whether the ‘computational average’ or the ‘high-level average’ hypotheses of ensemble coding is correct. In this study, we started by investigating the temporal ensemble coding of facial attractiveness. In a previous study, it has been shown that the visual system could involuntarily perceive the facial expressions via ensemble statistics (Ying & Xu, 2017). We adapted participants to a group of faces presented one at a time in RSVP (Potter, 1976). After prolonged exposure to a face for a few seconds the perception of its characteristics is diminished in subsequently viewed faces (Rhodes & Jeffery, 2006; Webster, Kaping, Mizokami, & Duhamel, 2004; Webster & MacLeod, 2011; Ying & Xu, 2017); so, adapting to an attractive face will lead to the subsequently viewed face as being less attractive (Pegors, Mattar, Bryan, & Epstein, 2015; Rhodes et al., 2003; Clifford, & Nakayama, 2003).

3.1. Experiment 2.1: Temporal ensemble statistics represent the computational average

The first experiment compared the adaptation aftereffects produced by the group of faces to those elicited by their computer-generated averaged face (the computational average): if they are indistinguishable from one another, then it would suggest that the brain is not extracting the gist, as many in the field believe, but instead computationally averaging all of the low-level properties of the faces together.

3.1.1. Methods

Participants

Twenty-nine participants (14 Females; Mean Age: 22.03) with normal or corrected-to-normal vision were recruited. In the beginning, we recruited 30 participants; however, one dropped out during the experiment. Thus, 29 participants finished this experiment. We selected this sample size based on previous research in facial attractiveness adaptation (n = 30 in Pegors et al., 2015). Written informed consent was provided by participants in all 3 experiments beforehand.

Procedure

Thirty-five Chinese female faces were chosen from the N-FEE database (Yap, Chan & Christopoulos, 2016). The N-FEE database contains high quality portrait pictures of Singaporean models, posting with various facial expressions. Due to copyright restrictions, we are not allowed to show these images in this thesis, so we have used faces from the KDEF dataset (Lundqvist et al., 1998) for demonstration. In this study, we only selected portrait pictures from 35 female Chinese Singaporeans with neutral expressions. All face images were grey scaled and masked so that only the facial region of each face was visible to the subjects. The luminance of the face images was equalized via SHINE toolbox (Willenbockel et al., 2010). Every participant rated the attractiveness of the 35 faces at least 2 weeks before the main experiment (adapted from Rhodes & Jeffery, 2006; 1 for most unattractive and 7 for most attractive): the adapting stimuli were selected from the four faces rated as most attractive and four that were least attractive. The test faces included one of the attractive and unattractive faces from the originally rated 35 faces (not the adaptors), and a further 5 faces that were produced by morphing these two faces in equal incremental steps between them (thus giving us 7 attractiveness units ranging from the original unattractive face through to the original attractive face) via Webmorph (DeBruine & Tiddeman, 2017). To minimize low-level adaptation as per prior research (Rhodes et al., 2003; Ying & Xu, 2017; Zhao & Chubb, 2001), the adapting stimuli were displayed at 3.20° × 4.03°, which is roughly 133% of the size of the test stimuli. The test stimuli were 2.40° × 3.02° in size, which is roughly 75% of the adapting stimuli. Similar to Haberman, Lee, and Whitney (2015), we are aware that the 'attractiveness unit' is arbitrary, and we do not mean that the (perceived) attractiveness differences between the testing faces are linear. The 'attractiveness unit' merely represents the physical difference between these faces.

Subjects completed 5 blocks: baseline, RSVP unattractive, RSVP attractive, computational average unattractive and computational average attractive. In the baseline condition, participants simply rated the test faces, which were presented for 400 ms, as attractive or unattractive. The temporal frequency of the RSVP sequence was 42.5 Hz, with each face displayed for 23.5 ms per face frame. Each test face was presented 10 times (a total of 70 trials). The same test face sampling occurred in the attractive RSVP block, but this time participants viewed an RSVP stream of the 4 attractive adapting faces prior to viewing each test face (Figure 3. displays the trial sequence). This method was repeated for the unattractive RSVP block, except the RSVP stream comprised the unattractive adaptors. The same process occurred for the attractive computational average block. Except during adaptation, participants were simply presented with a single face that was created by morphing all of the 4 attractive adaptors' low-level visual properties together. The same was true for the unattractive computational average block, except the unattractive adaptors were used to create its adapting face. The blocks were presented in randomized orders. Participants were given a break between blocks that were roughly equal in duration to an experimental block to disperse any carryover effects. Subjects practiced this and future experiments for 5-10 trials before data collection.

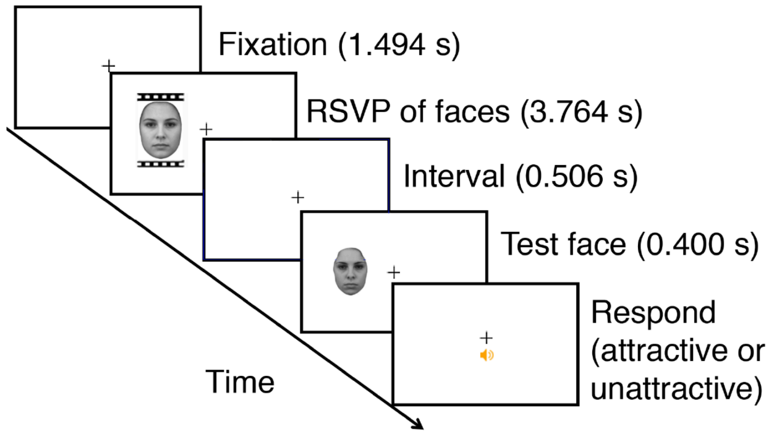

Figure 3.1. Example trial sequence from the RSVP adaptation condition (the demonstrated faces are AF01NES and AF34NES from the KDEF database). Participants fixated on the cross throughout the experiment. After 1.494 s, the RSVP of the faces appeared onscreen for 3.764 s. After a short interval (0.506 s), the test face appeared for 0.4 s. Then a beep sound prompted subjects to judge the target face by pressing the 'A' button as attractive, or the 'S' button as unattractive.

Analysis

Participants' responses were sorted into proportions of 'attractive' responses to each test stimulus per adaptation condition. A psychometric curve was created with the x-axis indexing the test stimuli and the y-axis plotting the fractions of 'attractive' responses. Calculation of PSE was the same as the previous study. We quantified the adaptation aftereffects by comparing the difference between the PSEs of the adapting conditions and that of the baseline condition. Any subsidiary pairwise comparisons after the analysis of variance (ANOVA) were Bonferroni corrected. To validate that the non-significant results truly supported the null hypothesis, ‘there was no effect’, we used the Bayes Factor to further analyze the data (Dienes, 2014; Rouder, Speckman, Sun, Morey, & Iverson, 2009; Morey, & Iverson, 2009). In brief, the Bayes Factor utilizes the observed evidence to compute the ratio between the likelihoods of the hypotheses. For instance, ‘$BF_{01} = 3$’ suggests that the observed result is 3 times more likely to fit the “null-hypothesis” compared to the “alternative hypothesis”. Generally, $BF_{01} > 3$ is recommended to provide evidence for the null hypothesis. All statistical analyses were conducted in JASP 0.8.6 (JASP team, 2018), R 3.4.3 (R Core Team, Vienna, Austria), Matlab R2017a (Mathworks, MA, USA) and SPSS Statistics 22 (IBM, NY, USA).

3.1.2. Results and Brief Discussion

To determine the presence of adaptation aftereffects in our experiment, we performed paired *t*-tests between the PSE of the baseline and the PSEs of the adaptation conditions (Figure 3.2). As expected, both the attractive RSVP and computational average conditions produced significant aftereffects (both ps $< .001$), with participants reporting the test faces as unattractive more frequently in the two attractive conditions relative to the baseline. Surprisingly, neither of the unattractive conditions produced any aftereffects (both ps $> .67$). Bayesian *t*-tests provided further support for the null hypothesis (both BF_{01}s > 4.65): the unattractive conditions did not generate significant aftereffects relative to baseline. Participants did not seem to be processing either set of unattractive adaptors as unattractive. This is contrary to the outcome predicted by the 'high-level top-down average' hypothesis, for if this hypothesis had been correct, then we should have observed aftereffects in the opposite direction to those found in our attractive conditions (i.e., negative relative to baseline, where test faces were rated as attractive more frequently after adaptation).

To test whether temporal ensemble perception was indistinguishable from the computational average (the Shapiro-Wilk tests indicated that all data follow normal distributions, *p*s > 0.75; except for 'Statica' condition, p = .001. Considering Ghasemi & Zahediasl's (2012) suggestion that with the current sample size '… we can use parametric procedures even when the data are not normally distributed'. We still applied parametric analysis here.), we performed a two-way repeated-measures ANOVA on the PSE shifts relative to baseline with factors of Attractiveness (attractive vs. unattractive) and Adaptor (RSVP vs. computational average). While there was a significant main effect of Attractiveness ($F(1,28) = 46.65$, $p < .001$, $\eta_p^2 = .0.63$) due to the attractive conditions having larger aftereffects than the unattractive conditions, there was no significant main effect of Adaptor ($F(1,28) = 0.028$, $p = .87$, $\eta_p^2 = .0.001$) nor any interaction ($F(1,28) = 0.345$, $p = .56$, $\eta_p^2 = .0.012$).

Bayesian *t*-tests comparing the attractive RSVP condition versus the attractive computational average (BF_{01} = 4.89), and the unattractive RSVP versus the unattractive computational average (BF_{01} = 4.32), provided further evidence for the null hypothesis. This confirms that the RSVP streams were processed by our participants in a similar way to their computational averages. Further support for this came from the fact that the attractive (r = .66, p < .001) and unattractive (r = .53, p = .003) RSVP streams were correlated with their computational counterparts.

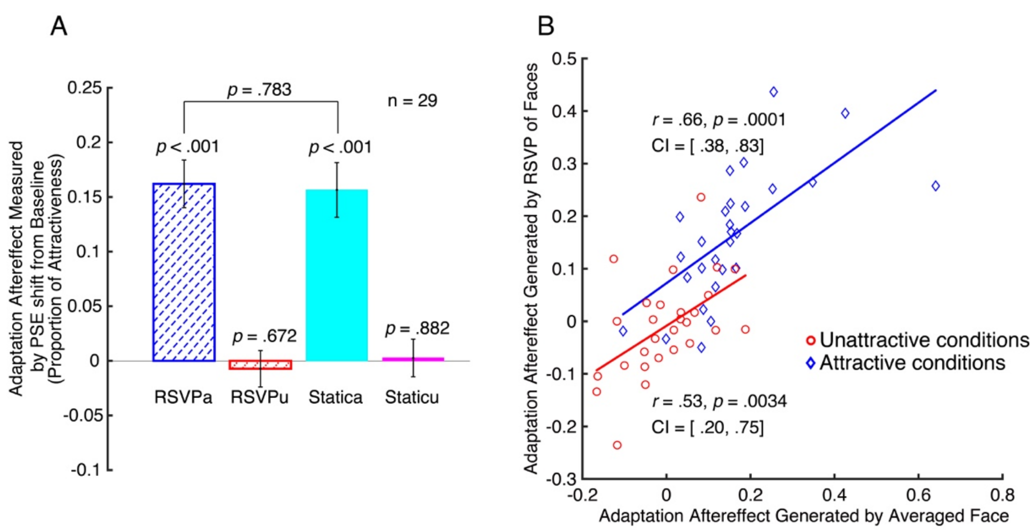

Figure 3.2. The RSVP and computational average aftereffects (Experiment 2.1). (A) Summary of all 29 subjects' results. RSVPa: adapting to attractive RSVP faces condition; RSVPu: adapting to unattractive RSVP

faces condition; Statica: adapting to the average face of the attractive RSVP faces condition; Staticu: adapting to the average face of the unattractive RSVP faces condition. For each condition, the adaptation aftereffect was measured by PSE shift relative to baseline (error bars were SEMs). The p-value above each bar was computed using paired t-tests. Noticeably, a positive adaptation aftereffect measured by PSE shift indicates the target faces were perceived as less attractive than during baseline. (B) The relationship between the RSVP conditions and the paired average conditions. Blue diamonds represent the attractive faces, and red circles represent the unattractive faces.

Adapting to an RSVP stream and its computational average led to comparable and correlated facial attractiveness aftereffects. These findings replicated prior work that showed similar effects on facial emotion (Ying & Xu, 2017, Chapter II). Moreover, the attractive and unattractive RSVP of faces generated asymmetrical adaptation aftereffects: unlike the attractive RSVPs, the RSVPs of unattractive faces which failed to generate significant adaptation aftereffects. Considering the similar and correlated aftereffects from the unattractive RSVP stream and the averaged face of them, it is reasonable to believe that the unattractive RSVP streams are represented as the ensemble average and were perceived as mediocre in facial attractiveness. Thus, the results here clarified what characteristics of a face our brain is extracting in order to produce temporal ensemble perception: contrary to the commonly used parlance in the literature (Haberman & Whitney, 2012; Whitney & Leib, 2017), temporal ensemble statistics do not actually represent the gist of the group's characteristics, but rather the computational average of the individual faces' low-level properties. This was clearly shown by the lack of differences between the RSVP streams' adaptation aftereffects and those produced by their computational averages. Had the brain been extracting the gist (i.e., that this group of faces is unattractive), then we should have seen differences in the adaptation aftereffects between the RSVP streams and their computational averages; an outcome that was not realized here. Similarly, the lack of aftereffects in the unattractive conditions relative to baseline further indicates that the unattractive group was not being processed as unattractive. Instead, the unattractive stream was indistinguishable from its computational average, which was more attractive than the underlying faces in the group, and thus comparable to participants' general levels of real-world habituation to facial attractiveness, as shown in the baseline condition. To our knowledge, this is the first time that any evidence has been produced to support the computational average hypothesis over the common 'gist' hypothesis.

3.2. Experiment 2.2: Temporal ensemble perception is linear

Based on previous work (Ying & Xu, 2017, Chapter II), it is anticipated that the brain will average the underlying mean attractiveness of an RSVP stream of faces in a linear fashion. This experiment, therefore, tested a separate group of participants with the two RSVP streams from Experiment 2.2, while including a 'Mixed' condition that comprised an RSVP stream of the attractive and unattractive adapting faces. If the brain averages the RSVP streams in a linear fashion, then there should be a positive correlation between the group's mean attractiveness ratings and its adaptation aftereffects.

In the Experiment 2.1, the data suggested that the RSVP of unattractive faces failed to generate significant adaptation aftereffects (asymmetrical aftereffects between attractive and unattractive RSVPs). Although we believe that this was caused by the computational mechanism of temporal ensemble coding of facial attractiveness. The data, however, is still open to an alternative explanation: our visual system might favor the attractive faces and ignore the unattractive ones (Palermo & Rhodes, 2007). If so, the insignificant aftereffect could be explained by the weaker cognitive resources. Thus, in this experiment, we added one new condition, the 'Mixed' condition, to elucidate this question. In this condition, all 4 attractive faces and all 4 unattractive faces used in the 'Attractive' and 'Unattractive' conditions were shown in the RSVP sequence. If the visual system just 'ignores' the unattractive faces and 'favors' the attractive ones, the four unattractive faces would be discounted in the ensemble representation, and the ensemble representation should be largely formed by the attractive faces. Thus, we should observe a significant adaptation aftereffect from the 'Mixed' condition, which is similar to that of the ‘Attractive’ condition. On the other hand, if our visual system is capable of encoding the unattractive faces (via the computational averaging, as we suggested in Experiment 2.1), then the attractive and unattractive faces would contribute (more or less) equally to the ensemble representation of the RSVP face stream. Thus, we should observe a significant adaptation aftereffect of the ‘Mixed’ condition which is similar to the mean of those of ‘Attractive’ and ‘Unattractive’ conditions. By monitoring the adaptation aftereffect of the new ‘Mixed’ condition, we could clarify the mechanism of the temporal ensemble coding of facial attractiveness.

3.2.1. Methods

Twenty new subjects (10 Females; Mean Age: 22.84) participated in this experiment. We selected this sample size basing on two reasons. Firstly, using the effect size of Experiment 1 ($\eta p2 = .65$), power analysis (using G*Power 3.1, with α-value at .05, and power ($1 - \beta$) at .80) indicated that we need at least 7 participants.

However, considering the differences in the experimental design, we chose to (roughly) triple the sample size.

We used the same procedure as in Experiment 2.1, except there were three adaptation conditions in addition to the baseline: RSVP of attractive faces (4 attractive faces), RSVP of mixed faces (4 attractive faces and 4 unattractive faces, from both attractive and unattractive conditions), and RSVP of unattractive faces (4 unattractive faces). All adaptors were the same as in Experiment 2.1. Also, each test face in each block appeared 14 times in a random order. We added more trials to investigate the robustness of the finding further. Additionally, after the main experiment, we asked the subjects to rate the mean attractiveness of the RSVPs on a 7-point scale (1 for most unattractive and 7 for most attractive), with each stream being presented 10 times.

A conventional correlation analysis has the assumption that the data are from independent observations. However, our data consisted of repeated measures from three observations (i.e., an observation from each of the unattractive, mixed, and attractive conditions) for each individual participant. We therefore used the repeated measures correlation analysis (Bakdash & Marusich, 2017) to quantify the strength of the relationship between the attractiveness ratings of the faces and the adaptation aftereffects produced by those faces.

3.2.2. Results and Brief Discussion

Relative to baseline, significant aftereffects were generated by the RSVPs of attractive (Figure 3.3a, $M = .20$, $SEM = .003$; $t(19) = 5.93$, $p < .001$) and mixed ($M = .11$, $SEM = .002$; $t(19) = 5.01$, $p < .001$) but not the unattractive ($M = .01$, $SEM = .02$; $t(19) = .63$, $p = .54$) faces (the Kolmogorov-Smirnov tests indicate that all data follow normal distributions, ps > 0.14). Bayesian analyses suggested that the lack of aftereffects in the unattractive condition was in favor of the null hypothesis (BF_{01} =

3.613); i.e., no adaptation aftereffect relative to baseline. Participants therefore rated the test faces as less attractive after adapting to the attractive and mixed RSVP streams. Moreover, this experiment replicated Experiment 1 in showing no aftereffects to the unattractive group, suggesting participants were actually not processing the group as unattractive. An ANOVA yielded significant differences among all three adaptation conditions (with Greenhouse-Geisser Correction, $F(1.46,27.70) = 34.46$, $p < .001$, $\eta_p^2 = .0.65$): subsidiary comparisons showed significant differences between the attractive and unattractive ($t(19) = 6.58$, $p = .001$), attractive and mixed ($t(19) = 4.24$, $p < .001$), and mixed and unattractive ($t(19) = 5.84$, $p < .001$) conditions. We further compared the adaptation aftereffects of the 'Mixed' condition with the average of those of the 'Attractive' and 'Unattractive' conditions. The paired sample *t*-test suggested that there is no significant difference between these two ($t(19) = .45$, $p = .68$). These indicate a graded fashion in which participants were perceiving the attractiveness of the adapting streams.

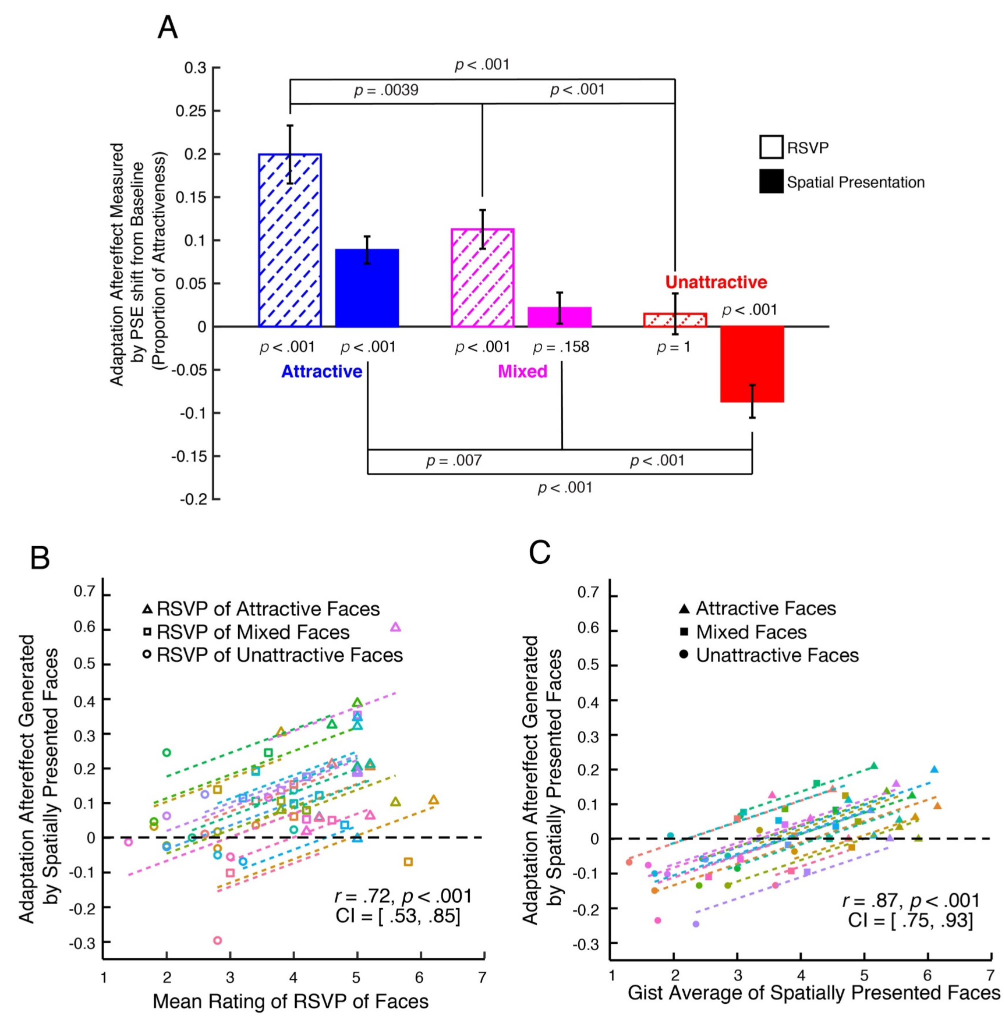

Figure 3.3. Adaptation aftereffects to RSVPs (Experiment 2.2) and spatially presented faces (Experiment 2.3). (A) Combined summary of all subjects' results of Experiment 2.2 and Experiment 2.3. The hatched bars indicate RSVP conditions (Experiment 2.2), and the solid bars represent Spatial Presentation conditions (Experiment 2.3). For each condition, the adaptation aftereffect was measured as a PSE shift away from the baseline (error bars were SEMs). The p-value shown above each bar was calculated by paired t-tests, comparing each adaptation condition with the baseline (no adaptation) or other adaptation condition in the same experiment.

Noticeably, a positive adaptation aftereffect measured by PSE shift indicates the test faces were perceived as less attractive than baseline. (B) The adaptation aftereffect as a function of the attractiveness rating of the RSVP of faces in Experiment 2.2. Hollow triangles represent the individual subjects mean ratings of RSVPs of attractive faces. Hollow squares for the RSVPs of mixed faces, and hollow circles for the RSVPs of unattractive faces. Each color represents one individual subject. The size of the adaptation aftereffect and the attractiveness ratings correlated significantly ($r = .72$, $p <. 001$). (C) The adaptation aftereffect as a function of the mean attractiveness rating of the adapting faces in Experiment 2.3. Filled triangles represented the individual subjects mean ratings of spatially presented attractive faces. Filled squares for the spatially presented mixed faces and filled circles for the spatially presented unattractive faces. Each color represents one individual subject. The size of the adaptation aftereffect and the attractiveness ratings correlated significantly ($r = .87$, $p < .001$). The faint dashed black auxiliary line indicates the no adaptation aftereffect.

An ANOVA on the subjects' attractiveness ratings of the RSVP streams suggested there were also significantly different from one another ($F(2,38) = 112.55$, $p < .001$, $\eta_p^2 = .0.86$). Further comparisons indicated that subjects judged the RSVP of the attractive faces ($M = 4.89$, $SEM = .013$) as the most attractive, followed by the RSVP of mixed faces ($M = 3.98$, $SEM = .016$), and the RSVP of unattractive faces ($M = 2.65$, $SEM = .017$) as least attractive (all $ps < .001$). A further repeated measures correlation analysis (Bakdash & Marusich, 2017) revealed a significant positive correlation between the RSVP streams' attractiveness ratings and their adaptation aftereffects (Figure 3.3b, $r = .72$, $p < .001$, 95% CI [0.53, 0.85]); indicating that the brain performs temporal ensemble statistics in a linear fashion from the underlying attractiveness of the stream.

In this experiment, we replicated the results from Experiment 2.1, but further illustrated the linear fashion in which the brain computationally averages the attractiveness of a temporal stream of faces. Moreover, the magnitude of the adaptation aftereffects of the 'Mixed' condition further supported our notion that the brain could perceive the unattractiveness of the unattractive faces. Thus, due to the mechanism of the temporal ensemble coding, the ensemble representation (reflected by the aftereffects) of unattractive faces could not alter the perception of the testing faces. These results, therefore, lend further support to our computational average hypothesis.

3.3. Experiment 2.3: Spatial ensemble statistics represent the gist

Across Experiments 1 and 2 we have shown that the temporal ensemble perception reflects the low-level bottom-up averaging. However, is this also true for spatial ensemble coding when a group of faces is presented simultaneously? To test this, we used the same groups of adapting faces as in Experiment 2, but presented them with a spatial configuration (Figure 4); using identical adaptor faces allows us to directly compare the aftereffects derived from temporal and spatial ensemble coding. If the aftereffects between Experiment 2 and 3 are indistinguishable, then it would imply that a similar mechanism is at work both temporally and spatially; i.e., the faces are being computationally averaged from their low-level properties. However, if the aftereffects between the two experiments are different, then it would provide the first evidence that temporal and spatial ensemble statistics actually reflect distinct mechanisms. If the 'high-level top-down average' hypothesis is at work during spatial ensemble coding, then we would expect an overall negative shift for all of the adapting face conditions: e.g., the unattractive group would elicit negative aftereffects, the mixed group would be no different from baseline, and the attractive group would elicit smaller positive aftereffects than the attractive group in Experiment 2.

3.3.1. Methods

Eighteen new subjects (11 Females; Mean Age: 22.78) participated in this experiment. We selected 20 subjects at the beginning (the same sample size as Experiment 2); however, 2 dropped out during the experiment. Here we used the same adapting faces and blocks from Experiment 3.2, except the mixed condition only contained 2 attractive and 2 unattractive. During adaptation, the 4 adapting faces were presented around the central fixation cross (Figure 3.4), with the test face presented in the center of the screen. The trial procedure was otherwise similar to Experiment 2.1 & 2.2. We also asked the subjects to rate the attractiveness of the eight individual adapting faces in a similar fashion as was performed in Experiment 2.1, and computed an average from the ratings reflecting the gist.

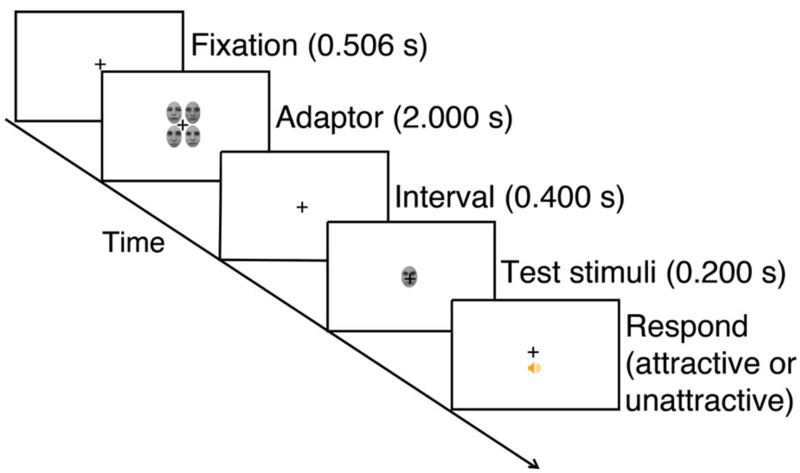

Figure 3.4. Example trial sequence from a spatial adaptation condition (the demonstrated faces are AF01NES, AF05NES, AF06NES, AF07NES and AF34NES from KDEF database). Subjects fixated on the cross at all times. After 0.5 s, four adapting faces simultaneously appeared for 2 s. After a 0.4 s interval, the test face appeared on the screen for 0.2 s. Then a beep sound indicated subjects should judge the attractiveness of the target face by pressing the 'A' button for attractive, or the 'S' button for unattractive. Experimental details can be found in the Methods section.

3.3.2. Results and Brief Discussion

Significant aftereffects were generated by both the attractive (*M* = .093, *SEM* = .016; *t*(17) = 5.91, $p < .001$) and unattractive (*M* = -.083, *SEM* = .020; *t*(17) = -4.20, $p = .001$) groups; test faces were rated unattractive or attractive more frequently relative to the no adaptation baseline following the attractive and unattractive adaptation groups respectively (the Shapiro-Wilk tests indicate that all data follow normal distributions, *p*s > 0.30). By contrast, the mixed faces did not evoke any aftereffects (*M* = .028, *SEM* = .019; *t*(17) = 1.48, $p = .16$). These results, therefore, show qualitative differences between the aftereffects of our previous two RSVP experiments and the spatial aftereffects here. An ANOVA on the three adaptation conditions was significant ($F(2,34) = 50.40$, $p < .001$, $\eta_p^2 = .0.75$). Bonferroni corrected comparisons indicated that the attractive and unattractive (*t*(17) = 8.68, *p* <.001), attractive and mixed (*t*(17) = 3.55, *p* = . 007), and mixed and unattractive

($t(17) = 7.93$, $p < .001$) conditions were all significantly different from one another. As with the other 2 experiments, there was a significant positive repeated measures correlation ($r = .87$, $p < .001$, 95% CI [0.75, 0.93]; Figure 3.3C) between the mean attractiveness ratings of the groups of adapting faces and their aftereffects.

To further clarify that the adaptation aftereffects produced by spatial versus temporal ensemble statistics were qualitatively different from one another, we performed a mixed model ANOVA on the adaptation aftereffects with a between participant factor of Group (Experiment 3.2: Temporal vs. Experiment 3.3: Spatial) and a within participants' factor of Attractiveness (unattractive vs. mixed vs. attractive). There was a significant main effect of Attractiveness (with Greenhouse-Geisser Correction, $F(1.56,56.19) = 78.75$, $p < .001$, $\eta_p^2 = .69$) due to differences between the adaptation aftereffects (i.e., attractive>mixed>unattractive, Figure 3.3, all $ps < .001$). Similarly, there was also a significant effect of Group ($F(1,36) = 11.20$, $p = .002$, $\eta_p^2 = .24$) due to the Experiment 2 Temporal group exhibiting more positive aftereffects in contrast to our current Spatial group [Exp 2 M = .11 vs. Exp 3 M = .012]. Finally, the Group × Attractiveness was not significant [with Greenhouse-Geisser Correction, $F(1.56,56.19) = .29$, $p = .70$, $\eta_p^2 = .0.08$]. These findings, therefore, indicate that while our participants were producing aftereffects that were similarly distinct between attractiveness blocks, the actual perceptual outcomes were qualitatively different.

It is obvious that the attractive and unattractive faces generated asymmetrical aftereffects in Experiment 2.2, while the very same faces generated symmetrical aftereffects in Experiment 2.3 (Figure 3.3A). To test whether the differences in aftereffects between the spatial and temporal aftereffects were simply due to the groups of participants perceiving the underlying faces' attractiveness differently, a mixed model ANOVA was performed on the mean attractiveness ratings with a between participant factor of Group (Temporal Experiment 2.2 vs. Spatial Experiment 2.3) and a within participants' factor of Attractiveness (unattractive vs.

mixed vs. attractive). There was a significant main effect of Attractiveness (F(2,72) = 302.74, p < .001, η_p^2 = .89) due to the faces being rated significantly differently from one another (i.e., attractive > mixed > unattractive, all ps < .001), but no main effect of Group (F(1,36) = .025, p = .88, η_p^2 = .01). There was, however, a significant interaction between the effects of Attractiveness and the Group (F(2,72) = 3.64, p= .031, η_p^2 = .092). However, while simple effects revealed significant within-group differences between all attractiveness levels (all ps < .001), there were no significant between-group differences in the mean attractiveness ratings of the adapting faces for each of the attractiveness blocks (attractive p = .11, mixed p = .82, unattractive p = .43). Thus, presenting the adapting faces spatially or temporally (RSVP) does not change their perception of the adapting faces' attractiveness *per se*, but simply their adaptation aftereffects. Therefore, these results suggest the qualitative differences in aftereffects between the RSVP and spatial experiments are not due to between-group differences in the perception of the adapting faces. Instead, it seems that spatial and temporal ensemble processing simply produce qualitatively distinct perceptual outcomes.

3.4. Discussion

This study investigated the perceptual operations underlying ensemble statistics across three experiments. Experiment 2.1 showed that RSVP streams and their paired computational averages created from the individual faces' low-level deviations led to comparable and correlated facial attractiveness aftereffects. Experiment 2.2 replicated the findings from Experiment 2.1 thereby further supporting the 'computational average' hypothesis. Moreover, Experiment 2.2 suggested that aftereffects increased as a function of the underlying RSVP stream's attractiveness. Incongruent with first 2 experiments, Experiment 2.3 showed that spatial ensemble statistics favored the gist hypothesis (i.e., no aftereffects in the mixed condition, negative aftereffects in the unattractive condition). Taking all three experiments together, it is clear that temporal and spatial ensemble statistics stem from qualitatively different extraction mechanisms.

While a number of prior studies have examined spatial ensemble coding and temporal ensemble coding (Haberman, Brady, & Alvarez, 2015; Haberman & Whitney, 2007, 2009; Whitney & Levi, 2011; Whitney & Leib, 2017; Wolfe, Kosovicheva, Leib, Wood, & Whitney, 2015; Ying & Xu, 2017), no study so far has compared the effects of both. Moreover, even if researchers had compared the averaging of facial traits other than attractiveness (e.g., emotion) across these two presentation formats, then they would have been unlikely to observe differences between temporal and spatial ensemble coding anyway. This is because adapting to facial emotion via either a low-level computational process, or a high-level averaging process, would result in the same perceptual outcome (as illustrated in Figure 1.1A with hypothetical data). Here, we took advantage of the fact that averaging faces together from their low-level properties makes them more attractive (DeBruine et al., 2007; Perrett et al., 1994; as illustrated in Figure 1.1B with hypothetical data). We showed that while both spatial and temporal ensemble statistics were driven by the attractiveness of the underlying faces, they resulted in different perceptual effects. Despite minor differences between the experiments, we do not believe that these differences can account for the diverging experimental outcomes. For example, in Experiment 3.1 & 3.2 the adapting faces were shown in the horizontal visual periphery (Ying & Xu, 2017), while in Experiment 2.3 the adapting stimuli were shown in a matrix around the center. However, we found no differences in how the two groups appraised their attractiveness, therefore indicating that the between-group differences in attractiveness aftereffects were not simply due to alterations in the underlying attractiveness of the adapting faces.

The other main difference between these spatial and temporal paradigms was that the test faces in Experiment 2.3 did not fully overlap with the adaptation faces. By contrast, the adaptation faces of Experiment 2.1& 2.2 were shown in the same screen location as the test faces. Typically, spatially overlapping test and adaptation faces produce massive aftereffects (Leopold, O'Toole, Vetter, & Blanz, 2001; Webster & MacLeod, 2011). If this could explain the differences in aftereffects between our experiments, then we would have expected the unattractive faces in the first two experiments to be more negative (or even more negative than that of Experiment 2.3) than those produced here. Similarly, the mixed conditions in Experiment 2.2 should have produced no aftereffects. However, our RSVP mixed condition produced positive aftereffects, reflective of the fact that these faces were processed by the brain as more attractive than their underlying mean compared to the gist average. The current data therefore clearly fit with the hypothesis that there are simply different types of ensemble statistics arising to produce qualitative between-group differences in aftereffects.

We should explicitly clarify to the reader that the null results found in

Experiments 2.1 and 2.2 (e.g., the unattractiveness conditions of Experiment 2.1& 2.2) actually support our low-level bottom-up averaging hypothesis. These findings are not caused by (a) insufficient unattractiveness in the unattractive adaptors to produce aftereffects, (b) nor by lacking power, (c) nor by ignorance of unattractive faces. First, we used the very same stimuli in Experiments 2.2 & 2.3, with the temporal and spatial presentation methods being the main difference between the two studies. The negative aftereffects generated by the unattractive condition in Experiment 2.3 shows that the unattractive group was being processed by the participants as unattractive; i.e., the participants perceived the subsequently presented faces as attractive. In the other words, the very same faces generated asymmetrical aftereffects when being presented *temporally*, but generated symmetrical aftereffects when being presented *spatially*. These findings are at odds with the suggestion that the unattractive adapting faces in Experiments 2.1 and 2.2 were simply insufficient in unattractiveness to elicit the aftereffects expected from an unattractive group. Such a suggestion is further strengthened by the large effect size in the unattractive condition's aftereffects in Experiment 2.3. Moreover, by analyzing the data via Bayes Factors (Dienes, 2014; Rouder et al., 2009), we found evidence supporting the null hypothesis (i.e., the unattractive RSVP faces are equivalent to baseline and their computer-generated average face), thus countering any suggestion that the null effects across Experiments 2.1 and 2.2 are a result of low statistical power. Also, the magnitude of the adaptation aftereffects observed in the 'Mixed' condition in Experiment 2.2 (see the result section of Experiment 2.2) clearly suggest that our vision system is capable of averaging the unattractive and attractive faces together; thus, our visual system does not ignore the unattractive faces. Simply put, the current data favors the notion that the RSVP streams of unattractive faces are perceived as neither attractive nor unattractive relative to participants' baseline norms of attractiveness, and that this perceptual outcome is not due to these faces being unattractive enough to elicit negative aftereffects. Instead, participants must have been averaging the RSVP stream in such a fashion that it made the faces as a group being processed as more attractive than their underlying mean. This was clarified by the fact that the aftereffects of the RSVP streams matched, and were correlated with, their more attractive computationally averaged morph face counterparts. In the future research, eye tracking measurement or fMRI is necessary to fully clarify the concerns

The qualitative differences between the adaptation aftereffects produced by RSVP streams and spatial presentations of faces likely reveal the hierarchical nature of the human face perception system (Bartolomeo, Vuilleumier, & Behrmann, 2015; Behrmann & Plaut, 2013; Duchaine & Yovel, 2015; Eimer, 2000; Gobbini & Haxby, 2007; Haxby et al., 2000, 2002; Liu, Harris, & Kanwisher, 2002; Young & Bruce, 2011; Zhao, Zhen, Liu, Song, & Liu, 2017). For example, extracting the computational properties of a face arguably occurs at an earlier stage of encoding

(Eimer, 2000; Gauthier et al., 2000; Grill-Spector, Knouf, & Kanwisher, 2004; Kanwisher, McDermott, & Chun, 1997; Kanwisher & Yovel, 2006; Pitcher, Walsh, Yovel, & Duchaine, 2007) in comparison to when the brain can conceptually calculate the aspects of a face that make it unattractive (O'Doherty et al., 2003). These factors probably reflect computationally averaged low-level qualities of stimuli that are extracted during temporal ensemble statistics versus the high-level general representation that is extracted during spatial ensemble statistics. Where these changes occur in the brain is at present unclear, but current data favor the hypothesis that temporal ensemble statistics likely arise from the early visual cortex, or early stage face areas of the brain, where low-level visual processing occurs continuously (Eimer, 2000; Haxby et al., 2000; Keysers, Xiao, Foldiak, & Perrett, 2001; Liu et al., 2002). By contrast, the FFA is likely to extract the high-level attractiveness of the face's structure, where a specific group of face-selective neurons is activated (Kanwisher et al., 1997; O'Doherty et al., 2003) prior to producing the gist. Support for this hypothesis stems from work showing an inverse relationship between neural responsiveness in the early visual cortex versus that produced by the fusiform gyrus when participants viewed serially presented faces (McKeeff, Remus, & Tong, 2007). For instance, activity in the early visual cortex became enhanced as serial presentations of faces become faster. However, such increasing of presentation speeds also led to a converse reduction in the fusiform gyrus's activity.

In conclusion, this study has shown for the first time that temporal ensemble statistics does not simply reflect the gist of a group of faces, but is instead extracted from the group's low-level computational average. Contrary to what many in the literature take as fact regarding the gist being extracted during ensemble perception, we have actually shown two distinct levels of ensemble statistics that can occur for the same facial trait: 'Gist averaging' for spatial ensemble coding and 'computational averaging' for temporal ensemble coding. These levels must reflect distinct neural encoding stages of the properties that make up facial attractiveness.

Chapter IV: Study 3: Ensemble Statistics Shapes Face Adaptation and the Cheerleader Effect

One of the widely studied and most important facial traits is that of attractiveness, due to its social and evolutionary significance (Little et al., 2011; Rhodes, 2006; Thornhill & Gangestad, 1999; Willis & Todorov, 2006). Researchers have uncovered numerous facial attributes that modulate facial attractiveness (Little et al., 2011; Rhodes, 2006). However, in addition to intrinsic qualities within a face, external factors such as context and experience may shape the perception of facial attractiveness (Anderson et al., 1973; Burriss, & Feinberg, 2007; Ewing et al., 2010; Hönekopp, 2006; Jones, DeBruine, Little, & Feinberg, 2007; Little et al., 2011; Penton-Voak et al., 2001; Rhodes, 2006). For example, simply being viewed in the company of others makes us appear more attractive than when we are seen alone; a powerful visual phenomenon known as the 'cheerleader effect' (Walker & Vul, 2014).

As is was put to us that the label ‘cheerleader’ effect might be a little outmoded and misleading, we henceforth refer to it as the ‘friend effect’ instead. Similarly, recent exposure to an unattractive distorted face also makes subsequently presented faces appear more attractive: an attractiveness adaptation aftereffect (Rhodes, 2006).

Although the external factors of facial attractiveness are relatively less studied, they have important implications in real life. Because we are commonly surrounded by friends in social situations where we are being judged by a prospective partner. It is important to ask do we look better when we surround ourselves with attractive or unattractive friends? More specifically as to whether the mean attractiveness of a group of faces affects the attractiveness judgments of a target face. This is an important question to answer as we are commonly surrounded by friends in social situations where we are being judged by a prospective partner, i.e., in a room, party or in photos used on social media apps. Despite the friend effect and adaptation being two widely examined phenomena thought to shape attractiveness judgments, no prior work has actually tested what influence a group's attractiveness has on any given target face being judged.

There are a number of competing theories as to how a surrounding group of faces may influence our perceptions of an individual's attractiveness, with each predicting a distinct outcome. The first, postulated by Walker and Vul (2014), would be a 'basking in reflected glory' or averaging effect hypothesis. This hypothesis predicts that faces are more attractive when judged in a crowd because the visual system tends to average all of the faces (including the target face and the surrounding faces) together. Under this hypothesis, faces judged in an unattractive crowd should, therefore, be perceived as less attractive than those faces when judged in an attractive crowd. This is because the average attractiveness of an unattractive crowd is less than that of an attractive crowd, as opposed to the 'basking in reflected glory' situation in which the target face receives a benefit from being in the company of attractive friends (DeBruine et al., 2007; Perrett et al., 1994).

Alternatively, there might be a contrastive effect, similar to what occurs during face adaptation (Calder et al., 2008; Hsu & Young, 2007; Leopold et al., 2001; Pegors et al., 2015; Rhodes et al., 2003; Rhodes & Leopold, 2011; Webster et al., 2004). In this case, the visual system produces a new norm from the ensemble representation of the surrounding faces which then affects the target's attractiveness ratings in the opposite direction to that of the crowd. Under these circumstances, being surrounded by increasingly unattractive friends will lead to a target's beauty being more readily perceived. This contrastive effect is because the unattractive group is thought to habituate the neurons that are tuned to unattractive faces, leading subsequent faces to be perceived as more attractive (O'Doherty et al., 2003; Pegors et al., 2015; Rhodes et al., 2003; Webster & MacLeod, 2011).

While the averaging and contrastive effects would arguably modulate the friend effect in line with the group's mean attractiveness, the two effects' relationships with the group's attractiveness would be in opposite directions to one another. For instance, under the averaging account, increasingly attractive friends would increase the target's perceived attractiveness. This is because in the averaging theory, the target face becomes more attractive as it is biased towards the mean of the crowd when averaged into the group. Support for this suggestion comes from prior work that shows a Gabor patch's orientation will contribute towards the group's average, yet also be perceived as that average (Morgan, Chubb, & Solomon, 2008; Ross & Burr, 2008). Converse to this, the contrastive hypothesis predicts that increasing group attractiveness would make a target face less attractive. Under such circumstances, the target face is compared to the ensemble representation of the surrounding faces. This contrastive effect is similar to that arising during facial attractiveness adaptation (O'Doherty et al., 2003; Pegors et al., 2015; Rhodes et al., 2003).

Finally, the mere presence of a face in a group may be sufficient to cause some kind of socially positive effect. In this case, simply being in a group makes a target face more attractive irrespective of how the group looks. What might cause this effect is unclear, but it may be driven by some kind of high-level process that can attribute popularity to a face due to the simple fact that it is surrounded by other people. It is possible that such a social positive effect might co-occur with either an averaging or contrastive effect. If this were to be true, then we might expect the target face in a group to always be rated as more attractive than when viewed in isolation, however, the strength of this friend effect may vary in response to the surrounding group's mean attractiveness.

While any of the above seems possible, recent research into ensemble statistics may give a hint as what effect the group's mean attractiveness might have. Our visual system rapidly and involuntarily averages the heterogeneous information from a group of faces presented simultaneously or in sequence to obtain its gist (Haberman

& Whitney, 2007, 2009; Sweeny & Whitney, 2014; Whitney & Leib, 2017; Ying & Xu, 2017). It is thus reasonable to imagine that such ensemble perception may also occur for facial attractiveness when viewing a group of faces (Abbas & Duchaine, 2008; Brady & Alvarez, 2015; Haberman & Whitney, 2012). One would expect ensemble statistics to help the visual system create a new norm for attractiveness from the group's average. This norm would form an implicit template against which any face presented in the group could be judged (Ying & Xu, 2017). Under these circumstances, we should expect increasingly unattractive groups to bias our judgments towards finding the target face as more attractive. This is because someone should be judged more desirable in an unattractive crowd as he/she is representing the best available option based upon current experience. Conversely, the same individual will not be judged to be quite as attractive when in an attractive group due to the fact that they are not as desirable as other options available. Similarly, we would anticipate ensemble perception to shape face adaptation in a similar way, albeit this time our perceptions are based upon recent, rather than concurrent, experiences. The following experiments aim to confirm these hypotheses.

4.1. Experiment 3.1. Face adaptation shows that we look better if we appear after a group of unattractive friends

If we appear after a group of friends, would we appear more attractive or unattractive? Could attractiveness aftereffect occur through spatial ensemble statistics by group face adaptation? When we are exposed to a single unattractive face for a few seconds, subsequently presented faces appear more attractive, with an attractive face producing a converse effect: a powerful visual illusion known as an adaptation aftereffect (Leopold et al., 2001; Rhodes et al., 2003; Webster et al., 2004). Similar adaptation aftereffects have been shown when people adapt to facial emotion (i.e., viewing a sad face makes subsequent faces seem happier; Afraz & Cavanagh, 2008; Burns, Martin, Chan, & Xu, 2017; Luo, Burns, & Xu, 2017; Webster et al., 2004; Xu, Dayan, Lipkin, & Qian, 2008), and we are able to extract the gist of the emotion from a group of faces through ensemble coding (Haberman & Whitney, 2009; Ying & Xu, 2017). It is currently unclear, however, whether this attractiveness aftereffect can occur through spatial ensemble statistics by adapting to a group of faces. To date, there has been a surprising lack of ensemble adaptation experiments in which multiple adaptors are simultaneously presented, with ensemble adaptation to low-level dots size one of the rare studies examining such an effect (Corbett, Wurnitsch, Schwartz, & Whitney, 2012).

If attractiveness perception can be similarly shaped by our prior experiences with groups of faces as those observed with the dots, then we would expect ensemble representations to shape face adaptation. Under such circumstances, a group of faces' mean attractiveness should predict their adaptation aftereffects. We tested this hypothesis in Experiment 3.1 by adapting participants to groups of faces that varied in their mean attractiveness and asking them to make attractiveness judgments to subsequently presented faces.

4.1.1 Method

Subjects

Twenty ethnic Chinese subjects (11 females, mean age 22.95 years), with normal or corrected-to-normal vision, participated in both experiments. We selected this sample size because it is comparable to previous research examining the ensemble coding using face adaptation aftereffects (we doubled the sample size from Ying & Xu, 2017). From the post-hoc power analysis (with α-value of .05, $\eta_p^2 = .64$, G*Power 3.1), we found this sample size yield a high power $1 - \beta = 1.00$. All participants gave their written consent before the study.

Apparatus

The same as previous studies.

Stimuli

All of the visual stimuli were female faces chosen from the Oslo face database (Chelnokova et al., 2014). The Oslo face database contains roughly 200 male and female faces, who were students (mostly Caucasian) at University of Oslo, with neutral expressions. Due to copyright restrictions, we are not allowed to show these images in this thesis, so we have used faces from the KDEF dataset (Lundqvist et al., 1998) for demonstration. We chose this face database for two reasons. First, it contained high-quality pictures that varied in attractiveness. Second, judgments of attractiveness towards female faces are almost perfectly correlated irrespective of the race being judged or judging (i.e., $r > .9$ in Cunningham, Roberts, Barbee, Druen, & Wu, 1995; Perrett et al., 1998; Rhodes et al., 2001). Thus, we anticipated that our Chinese Singaporean participants would have little difficulty in processing the attractiveness of the Caucasian faces in a normal manner. All face images were grey scaled and had an oval shaped crop applied so that only the central region of each face was visible by Paint.Net (dotPDN LLC, USA) and Matlab R2010a (Mathworks, MA, USA).

We assessed the perceived attractiveness of the stimuli via an online pilot study. Twenty subjects, those who took part in our two main experiments, were asked to rate the facial attractiveness for all 30 faces. Each face's attractiveness was assessed on a 7-point scale (1 for most unattractive and 7 for most attractive). All faces were presented in a random order, with this cycle being repeated 3 times. Each time a face was presented, it would remain onscreen until a judgment was made before starting the next trial. The mean attractiveness ratings for each face ranged from 2.10 to 5.28 (M = 3.52, SD = .85). The inter-rater reliability was high (Cronbach's alpha = .94) as has been shown in prior work examining attractiveness judgments (DeBruine et al., 2007; Rhodes et al., 2001). We selected the most attractive (M = 4.93, SD = .24) and most unattractive (M = 2.60, SD = .22) faces identified by our participants as adapting and test faces based on these ratings. Due to publishing restrictions, we use faces from Karolinska Directed Emotional Faces database in our figures (Lundqvist, Flykt, & Öhman, 1998).

Test Faces. Using MorphMan 2016 (STOIK Imaging, Moscow, Russia), we morphed the top two attractive faces together to create our most attractive target/test face. We did this as we wanted to make our highly attractive face even more attractive, and face averaging achieves this goal as average faces are rated more attractive than their non-average counterparts (DeBruine et al., 2007; Perrett et al., 1994). This face was then morphed with the most unattractive face to create a sequence of seven, incrementally spaced, morph continua test faces; we did not average the most unattractive face as we wanted to have the most unattractive face possible. The morphed faces were separated by units in proportions of $1/7^{th}$ attractiveness. For example, the most attractive morph which contained 100% of the attractive face (and 0% of the unattractive face) was equal to 1 attractiveness unit; the least attractive morph with 0% of the synthesized attractive face and 100% of the unattractive face was equal to 0 attractiveness units. All test stimuli subtended a horizontal and vertical visual angle of 1.80° × 2.40° respectively.

Adapting Faces. After selecting the most unattractive face from the Oslo face database for the test face, we selected the six most attractive faces and the six most unattractive faces from the remaining faces as the adapting stimuli. There were three types of adapting faces: all 6 most attractive faces; all 6 most unattractive faces; or a mixture of 3 most attractive and 3 most unattractive faces (randomly selected from the 6 attractive and 6 unattractive faces). These faces were displayed at the same size as the test faces.

Procedure

We used a block design comprising three experimental blocks and one baseline block. In the attractive adaptation block, the 6 presented adapting faces were the 6 most attractive faces. In the unattractive block, the faces were the most unattractive faces. In the mixed block, 3 of the attractive faces and 3 of the unattractive faces were presented. In the baseline condition block, only the test face was presented with no adapting faces. We favored a block design as it meant that participants were judging the test faces within a consistent group context within each block. In each trial, the test stimulus presented was one of the morphed faces selected at random. Each trial was initiated with a central fixation cross for 500 ms. This cross would be present in all trials and participants were requested to remain fixated at the cross at all times. The 6 adapting faces would then surround the central fixation cross in a hexagon fashion (See Figure 4.1) for 1 s. A 400 ms interstimulus interval would then occur with only the fixation cross present, before the test face's presentation, superimposed under the fixation cross, for 200 ms. Therefore, the test face was presented onscreen after the group of 6 adapting faces, as is usual in adaptation studies (Rhodes et al., 2003; Webster et al., 2004; Ying & Xu, 2017). There was no spatial overlap among the adapting and test faces, so any adaptation aftereffects arising would not be low-level retinotopic effects but instead require higher level adaptation (Rhodes et al., 2003). A final screen with only the fixation would then be presented until participants pressed the key response to indicate whether the test face was attractive or unattractive ('A' for attractive, 'S' for unattractive). This screen commenced with a 50 ms noise sound to alert participants to respond, with their response starting the next trial (Figure 4.2).

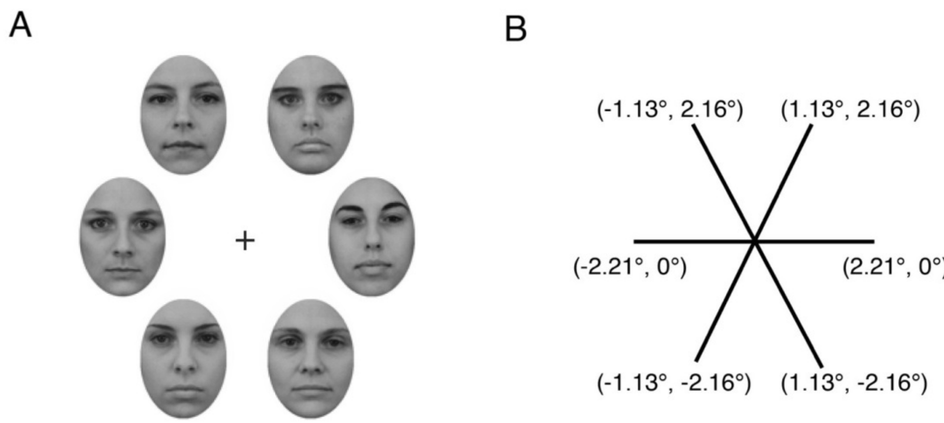

Figure 4.1. An example of the adapting faces used in Experiment 1 (the demonstrated faces are AF01NES, AF06NES, AF08NES, AF17NES, AF20NES, and AF34NES from KDEF database). (A) The 6 adapting faces formed a hexagon. In the experiment, the central fixation cross was right in the center of them. (B) The schematic of the relative locations of the stimuli. The test face was presented in the central position (the intersection of the three lines) of Cartesian coordinates. The locations of the central points of the adapting faces are at the endpoints of each line; the coordinates are the relevant location for each adapting face. For example, the top left surrounding face is 1.13° to the left and 2.16° above the fixation cross.

In each block, each test face was presented 10 times (in random sequences) making a total of 70 trials in each block. Similarly, the locations of the six adapting face identities were also shuffled randomly around the hexagon. Each block lasted around seven minutes, and there was a seven-minute rest in between consecutive blocks to avoid any carryover effects. The order of the blocks was randomized for each subject. Data collection initiated after sufficient practice trials.

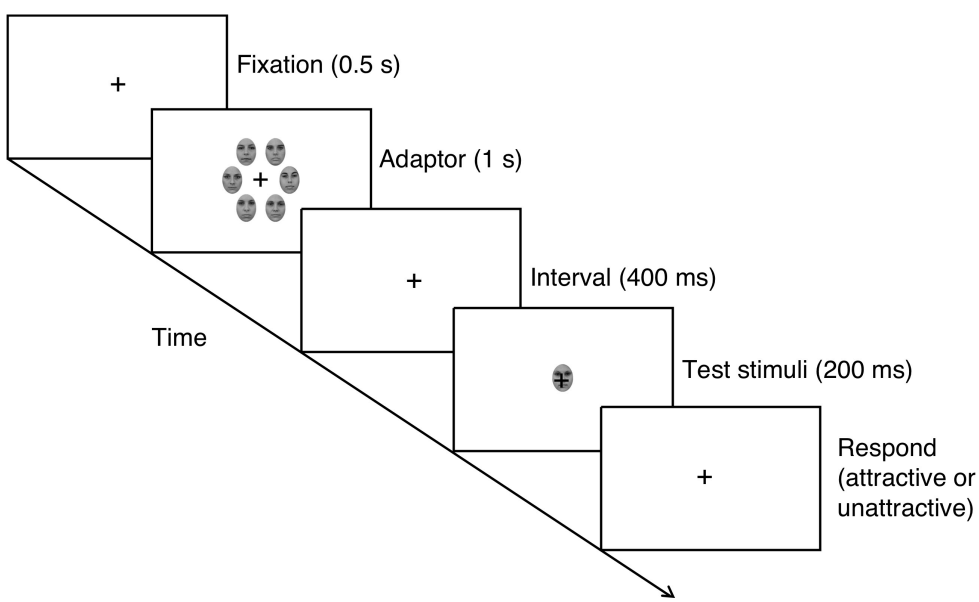

Figure 4.2. The sequence of an adaptation trial (the demonstrated faces are AF01NES, AF06NES, AF08NES, AF17NES, AF20NES, AF26NES, and AF34NES from KDEF database). Subjects pressed the space bar to start a block. After 500 ms, the adaptors, six faces appeared at the screen for 1 s. The locations of the adaptors are the same as the surrounding faces in Figure 4.1. Then after a 400 ms interval, the test face appeared at the center of

the screen for 200 ms. Subsequently, a beep sound prompted subjects to judge the attractiveness of the central face by pressing the 'A' button for attractive, or the 'S' button for unattractive.

Data analysis

The calculation of PSE was adapted from previous studies. The adaptation aftereffect (and friend effect in later experiments) was quantified as the difference between the PSE of each experimental condition relative to the baseline. We used repeated measures ANOVA and pairwise comparisons (Bonferroni corrected) to compare subjects' PSEs for different conditions. The means derived from the attractiveness of the adapting faces were calculated by averaging the ratings of each adapting face by each subject individually.

Considering the fact that all participants were repeatedly measured under three adaptation conditions, we used repeated measures correlation analysis (Bakdash & Marusich, 2017) to measure the correlation. The statistical analyses were conducted in R 3.4.3 (R Core Team, Vienna, Austria), Matlab R2010a (Mathworks, MA, USA) and SPSS Statistics 22 (IBM, NY, USA).

4.1.2. Results and Brief Discussion

The results from all participants averaged together are presented in Figure 4.3A. We plotted the proportion of "attractive" responses as a function of the "attractiveness" units of the test faces. Adaptation aftereffects can be interpreted from the psychometric curve shift. The black dash-dotted line psychometric curve is the baseline condition. After exposure to the unattractive faces (solid blue line), there was a leftward shift in the psychometric curve relative to baseline. A similar shift, albeit in the opposite direction, is present in the attractive group condition (magenta dotted line). Moreover, the mixed group, in which the attractive and unattractive faces appear to cancel each other out, shows no shift compared to the baseline (cyan dashed line). The psychometric curves in the attractive and unattractive conditions illustrate the existence of classic adaptation aftereffects (Webster & MacLeod, 2011;

Ying & Xu, 2017).

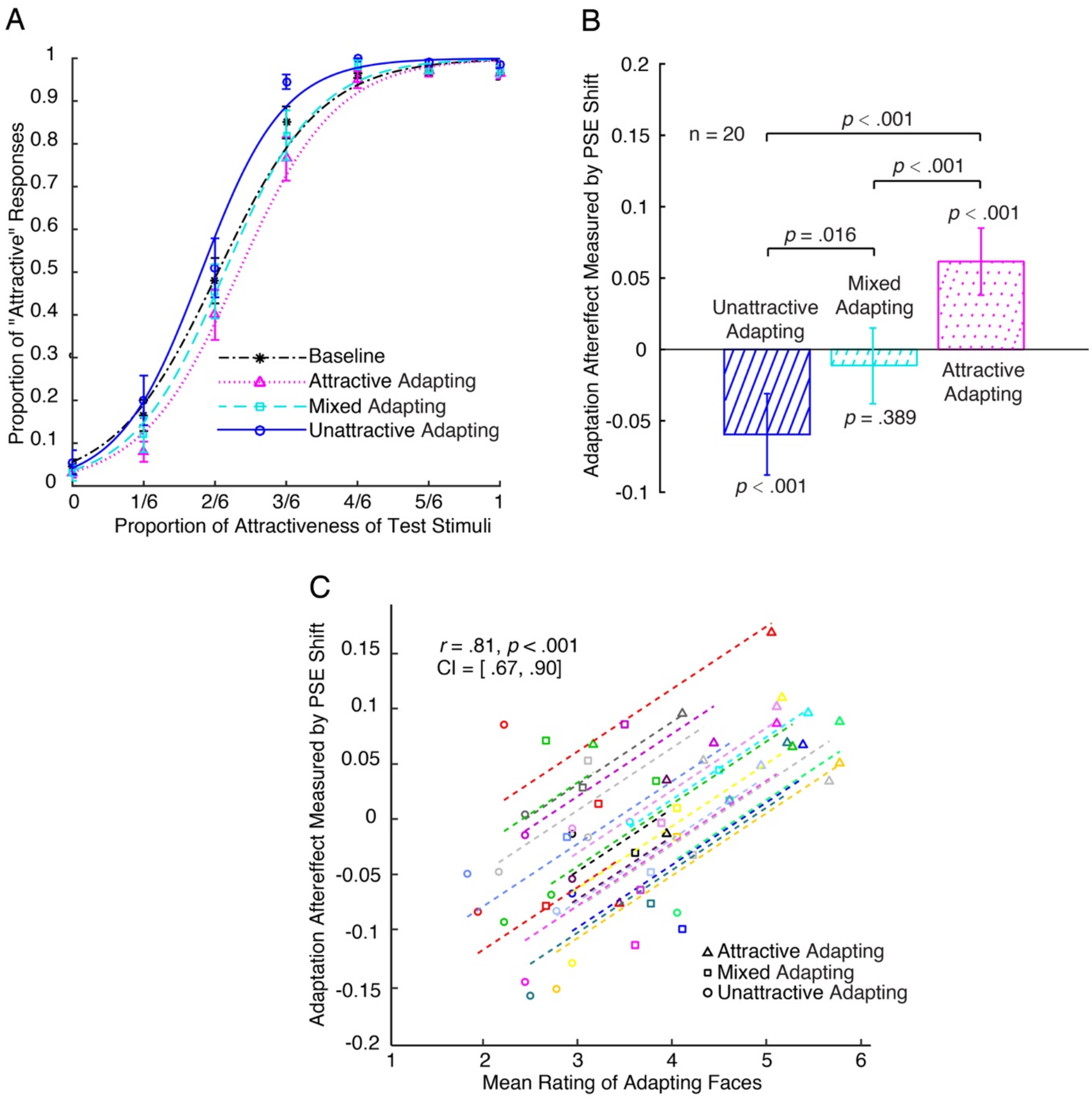

Figure 4.3. The attractiveness adaptation aftereffects of adapting faces with different levels of mean attractiveness (Experiment 3.1). (A) The psychometric functions of all subjects averaged together. 'Baseline' represents the baseline condition without any adapting faces (black star, black dash-dotted line). 'Attractive Adapting' represents the attractive adaptation condition with six attractive faces during adaptation (magenta

triangle, dotted line). Error bar indicates the standard error of the mean. 'Mixed Adapting' represents the mixed adaptation condition with three attractive and three unattractive faces during adaptation (cyan square, cyan dashed line). 'Unattractive Adapting' represents the unattractive adaptation condition with six unattractive faces during adaptation (blue circle, blue solid line). (B) Summary of all 20 subjects' results. For each condition, the average PSE relative to the baseline condition and the 95% confidence intervals were plotted. The p value shown for each condition in the figure was calculated against the baseline condition using two-tailed paired t tests. Note that a more negative adaptation aftereffect measured by PSE shift indicates that the test faces were perceived as more attractive than with no adaptation. (C) The attractiveness adaptation aftereffects as a function of the mean attractiveness ratings of the adapting faces. Each data point is derived from the mean attractiveness rating of the adapting faces and their aftereffect from a single observer for each adaptation condition. Thus, each participant has his/her own correlation line fitted to the data points, of the same color. Taken together, the size of adaptation aftereffects and the mean attractiveness ratings correlated significantly ($r = .81$, $p < .001$, 95% CI [0.67, 0.90]).

We then compared the mean PSEs relative to the baseline of all 20 subjects to quantify the facial attractiveness adaptation aftereffect (Figure 4.3B). The Shapiro-Wilk tests indicate that all data follow normal distributions, ps > 0.15. Positive values represent the rightward (less attractive judgment) shifts of the respective psychometric curves; and negative values represent the leftward (more attractive judgment) shifts of the respective psychometric curves. Paired t-tests revealed significant negative adaptation aftereffects in the unattractive (i.e., test faces were more likely to be judged as attractive; *M* = - 5.96%, *SEM* = .01; *t*(19) = - 4.33, *p* < .001, Cohen's *d* = 0.97, 95% CI [- 0.09, - 0.03]) and positive aftereffects in the attractive (i.e., test faces more likely to be judged as not attractive; *M* = 6.16%, *SEM* = .01; *t*(19) = 5.49, *p* < .001, Cohen's *d* = 1.23, 95% CI [0.04, 0.09]) conditions. By contrast, the mixed adapting faces yielded no significant aftereffects (*M* = - 1.12%, *SEM* = .01; *t*(19) = - .88, *p* = .388, Cohen's *d* = 0.20, 95% CI [- 0.04, 0.02]). A repeated measures ANOVA also indicated significant differences among all three adaptation conditions ($F(2,38) = 33.48$, $p < .001$, $\eta_p^2 = .64$). Further comparisons indicated significant differences between the unattractive and the attractive (*t*(19) = 7.63, *p* < .001, Cohen's *d* = 1.71) conditions, and between the unattractive and mixed conditions (*t*(19) = 3.16, *p* = .005, Cohen's *d* = .71). A significant difference was also found between the mixed and attractive adaptation conditions (*t*(19) = 5.43, *p* < .001, Cohen's *d* = 1.22).

To investigate if ensemble averaging shapes adaptation, we analyzed the repeated measures correlation between the adaptation aftereffects and the mean attractiveness ratings across all group conditions (Figure 4.3C). A significant association was revealed between these mean attractiveness ratings and the size of the aftereffects (r = .81, p < .001, 95% CI [0.67, 0.90]). Ensemble averaging, therefore, appears to drive the creation of a new attractiveness norm which then acts as a virtual adapting stimulus.

4.2. Experiment 3.2. The adaptation aftereffect is subject to the ensemble perception of the adaptors

Experiment 3.1 showed that a group of faces could produce adaptation aftereffects. We wondered if this occurred from an averaging or summation process. In Experiment 3.2 we tested these possibilities in a number of different ways. First, we generated a morphed average face (Figure 4.4B) from the attractive face group (Figure 4.4A) and examined whether it could generate similar attractiveness aftereffects; if ensemble averaging was occurring, then the attractive group should produce adaptation aftereffects that are equal to their morphed average group, as the means of both groups are equal.

The second way we tested summation versus averaging was to assess whether ensemble adaptation led to distinct aftereffects when compared to the processing of an individual face (Figure 4.4C) from the group (Figure 4.4A). If summation was occurring, then the adaptation aftereffects produced by the single face should be roughly equivalent to 1/6th of the attractive group of faces. Similarly, the aftereffect by adapting to a group of the same single face presented in six locations at the same time (Figure 4.4D) should be larger than that produced by the single face (Figure 4.4C). Please note, our reason for picking only the attractive faces was simply due to the impractical time constraints of testing all possible permutations from Experiment 3.1. And also, we are aware that the attention distribution might be different between the 'Single1' and 'Single6' conditions; however, testing the 'Single1' condition could offer us a unique chance to check whether the adaptation aftereffect of the ‘Single6’ condition is just a summation of the adaptation aftereffects of each individual adapting faces or not.

4.2.1. Methods

Subjects, Apparatus, Stimuli, and Procedure

Thirty new participants took part in this experiment (two of them are authors). We chose this sample size for two reasons. The result of power analysis (using G*Power 3.1 software, basing on $\eta_p^2 = .64$ from Experiment 3.1, with α-value at .05, and power (1 – β) at .80) indicated that we need at least 6 participants. We further considered that recently Pegors and colleagues used 30 as the sample size (2015) in their facial attractiveness adaptation study. Therefore, we chose 30 as the sample size of the current experiment. We used the same lab setting, analysis, and the face dataset as in Experiment 3.1, except for a couple of changes.

Firstly, to rule out the possible confounding factor from the overlap between the test faces and adapting faces, we created a new stream of the testing faces and carefully selected adapting faces. In this experiment, under the same manipulation, the new testing faces are a morph continuous of the most attractive face and the least attractive face from the face database (based on Experiment 3.1). The adapting faces are the six most attractive faces and the six least attractive faces from the **rest** of the database. Therefore, the test faces and the adapting faces are from different identities. As in Experiment 3.1, the faces are cropped with an oval shape mask. Moreover, the luminance of the faces was further equalized by the SHINE toolbox (Willenbockel et al., 2010).

Secondly, as aforementioned, we have four different adaptation conditions (Figure 4). For the AVE condition, we created the averaged face of the adaptors using the Webmorph software (DeBruine & Tiddeman, 2017) to average all of the faces from the attractive group (Figure 4.4A). For the Single1 condition, we picked one of the faces (note that, this face is not the most attractive one from the candidates; we picked it up randomly) from the attractive group adaptors and presented it at one of the 6 locations randomly from trial to trial. To match the low-level features with the ATT condition, we created scrambled faces from the rest of the attractive adaptors respectively via the Webmorph software (DeBruine & Tiddeman, 2017), and presented the scrambled faces in the other 5 adapting locations in the group. Finally, we created the Single6 group by simply presenting the Single1 face in all six adaptation locations (Figure 4.4D).

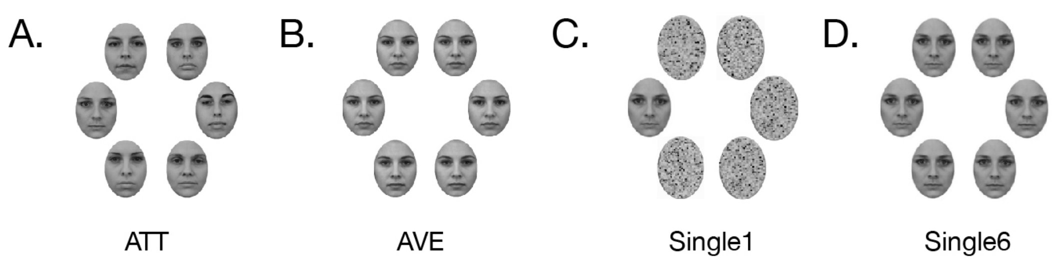

Figure 4.4. The adapting faces for Experiment 3.2 (the demonstrated faces are AF01NES, AF06NES, AF08NES, AF17NES, AF20NES, AF26NES, and the averaged face of them from KDEF database). (A) The attractive adaptors (ATT) condition. (B) The averaged face (AVE) condition, the faces are all the averaged face of the six attractive adapting faces. (C) The single face with scrambled faces (Single1) condition. (D) The single face repeated six-time (Single 6) condition.

4.2.2. Results and Brief Discussion

The results from all participants averaged together are presented in Figure 4.5A. We plotted the fraction of 'attractive' responses as a function of the attractiveness units of the test faces. The adaptation aftereffect can be interpreted from the psychometric curve shift: the leftward shift means the test faces are perceived as more attractive, and the rightward shift means the test faces are perceived as less attractive. All four conditions generated significant rightward shifts. Moreover, there is no significant difference between the PSE shifts of the ATT and AVE pair, as well as the Single1 and Single6 pair.

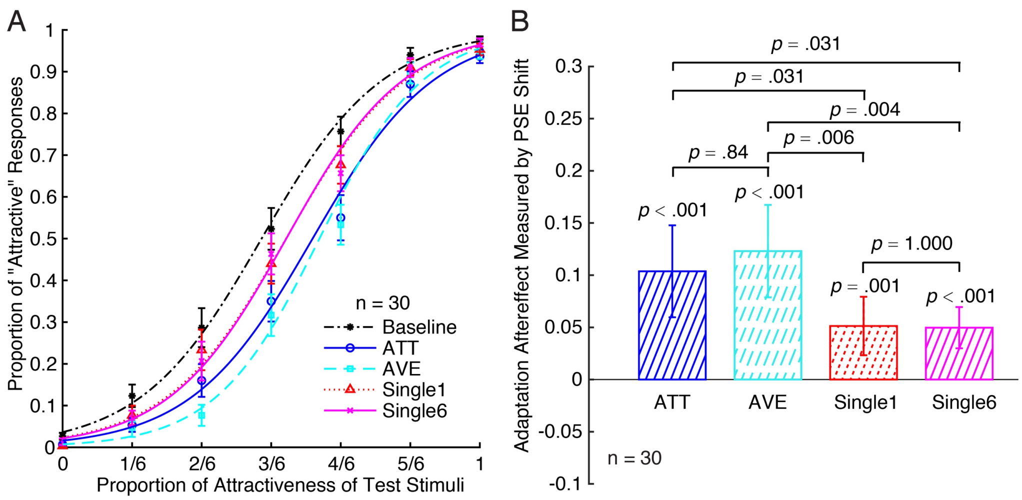

Figure 4.5. The attractiveness adaptation aftereffects of surrounding faces with different levels of mean attractiveness (Experiment 3.2). (A) The psychometric functions of all participants averaged together. 'Baseline' represents the baseline condition without any adapting faces (black star, black dash-dotted line). 'ATT' represents the attractive adapting faces condition with six attractive faces (blue circle, solid line). Error bar indicates the standard error of the mean. 'AVE' represents the AVE adaptation condition with six averaged faces of the ATT condition (cyan square, dashed line). 'Single1' represents the Single1 adaptation condition with one attractive face and the scrambled faces of the other five attractive faces (red triangle, dotted line). 'Single6' represents the Single6 adaptation condition with six repetitions of one attractive face during adaptation (magenta X, solid line). (B) Summary of all 30 participants' results. The *p*-value above each bar was calculated against the baseline condition using the pairwise comparison (Bonferroni corrected). Note that, a more positive adaptation aftereffect measured by PSE shift indicates that the target faces were perceived as less attractive than on their own.

The summary of the adaptation aftereffects measured by PSE shift is illustrated in Figure 4.5B. The Shapiro-Wilk tests indicate that all data but those from AVE condition follow normal distributions, *p*s > 0.15. Ghasemi & Zahediasl (2012) suggested that with the current sample size '… we can use parametric procedures even when the data are not normally distributed'. Therefore, we still applied parametric analysis in here. Compared to the baseline PSE, the ATT (M = 10.37%, SEM = .022; $t(29) = 4.81$, $p < .001$, Cohen's d = 0.88, 95% CI [0.06, 0.15]) , AVE (M = 12.30%, SEM = .022; $t(29) = 5.69$ $p < .001$, Cohen's d = 1.04, 95% CI [0.08, 0.17]), Single1 (M = 5.13%, SEM = .014; $t(29) =$ 3.74, $p = .001$, Cohen's d = 0.43, 95% CI [0.2, 0.8]), and Single6 (M = 5.00%, SEM = .010; $t(29) =$ 5.15, $p < .001$, Cohen's d = 0.42, 95% CI [0.3, 0.7]) conditions all generate significant adaptation aftereffects. A repeated measures analysis of variance (ANOVA) also indicated significant differences among all four adaptation conditions (Mauchly's test indicated that the assumption of sphericity was violated, $\chi^2(5) = 14.47$, $p = .013$; thus the degree of freedoms were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .73$); $F(2.18,63.13) = 9.72$, $p < .001$, $\eta_p^2 = .25$). Further Bonferroni corrected pairwise comparisons indicated that there were no significant differences between the ATT and AVE conditions ($t(29)=1.52$, $p = .84$, Cohen's $d = .28$), nor the Single1 and Single6 conditions ($t(29)=.124$, $p = 1.00$, Cohen's $d = .023$). Moreover, both ATT and AVE conditions generated significantly larger aftereffects than both Single1 and Single6 conditions (all *p*s < .031). Finally, the correlation analysis indicated that there were significant correlations between ATT and AVE conditions ($r = .83$, $p < .001$, 95% CI [0.66, 0.91]), as well as between Single1 and Single6 conditions ($r = .44$, $p = .015$, 95% CI [0.10, 0.70]). Noticeably, the adaptation aftereffect of the Single1 condition ($M = 5.13\%$) was much larger than 1/6 of those of the ATT and Single6 conditions (10.37% and 5.00%, respectively). In order to test whether the Single1 face's adaptation aftereffects were a $1/6^{th}$ of the ATT group's aftereffects, we performed a one-sample *t*-test on the Single1 condition's aftereffects, comparing to the $1/6^{th}$ of the ATT group's mean aftereffect value (M = 1.73%). The aftereffects in the Single1 condition were significantly larger than this value ($t(29)=2.41$, $p = .019$; Cohen's $d = 1.43$), hence indicating that the ATT group's aftereffects were unlikely to have arisen through summation.

In summary, our results from the first two experiments indicate that adaptation aftereffects can arise from a group of faces. Moreover, these effects do not appear to be the result of each individual face being adapted to and summed together, but instead, the aftereffects seem equal to those produced by their averaged counterparts (ATT and AVE conditions). Similarly, a single face (Single1) produced equivalent aftereffects to those resulting from a group of the same face (Single6). These findings taken together support the hypothesis that the brain averages the faces in a scene together to produce adaptation aftereffects. From this, we were therefore curious if ensembles of faces influenced another face perception phenomenon related to facial attractiveness: the ‘friend effect’. Specifically, how does this friend effect vary with the attractiveness of the surrounding faces?

4.3. Experiment 3.3 We look better with unattractive friends

In the previous two experiments, we found that ensemble statistics could influence face adaptation. We were therefore curious if ensemble perception could also influence another phenomenon related to face perception: the friend effect. The friend effect is characterized by an individual face being perceived as more attractive when it is viewed in the presence of other faces (‘friends’), in contrast to when it is judged in isolation by itself (Walker & Vul, 2014). We wanted to test whether ensemble perception could similarly modulate the magnitude of this friend effect as we had observed in our adaptation studies. Therefore, we asked participants to judge the facial attractiveness of a central target face when it was either presented by itself in a baseline condition or surrounded by a group of faces that were attractive, unattractive or mixed (the ‘friend’ conditions). By employing such a paradigm, we would be able to ascertain what, if any, influence ensemble perception was likely having on the friend effect. Moreover, by using the same faces between the first experiment and this one, we would be able to directly study whether there is any relationship between the friend effect and face adaptation.

There are a number of competing theories for how ensemble perception may influence the friend effect. The first, an averaging effect postulated by Walker and Vul (2014), predicts that faces are more attractive when judged in the crowd due to the crowd biasing the perception of that face towards the group's average, and thus more attractive. If this is the case, then faces judged in an unattractive crowd should be judged as less attractive than those faces when judged in an attractive crowd. This is because the average attractiveness of an unattractive crowd is less than that of an attractive crowd (DeBruine et al., 2007; Little, 2007; Perrett et al., 1994). Alternatively, the friend effect might be explained by a contrastive adaptation effect, as we saw in Experiment 3.1, whereby a new norm is created from the surrounding crowd which then influences the attractiveness ratings of the central face. Thus, being judged in an attractive crowd will make a face appear less attractive than when judged in an unattractive crowd. Under such circumstances, a group of faces' mean attractiveness should predict their adaptation aftereffects and friend effects. As these latter two are shaped by ensemble statistics, we would anticipate a significant relationship between them too. Finally, there may be a social positive effect, where the mere presence of 'friends' boosts your attractiveness irrespective of their looks. This latter effect may, however, also occur concurrently with the averaging or contrastive theories we described above. We examined these hypotheses in Experiment 3.3.

4.3.1. Methods

Subjects, Apparatus, Stimuli, and Procedure

The same 20 subjects from Experiment 3.1 participated in this experiment. From the post-hoc power analysis (with α-value of .05, $\eta_p^2 = .35$, G*Power 3.1), we found this sample size is sufficient to yield a high power: $1 - \beta = .98$. We used the same lab setting and stimuli as in Experiment 3.1, except for a couple of minor adjustments. The key difference was that in the present paradigm, the surrounding faces were presented at the same time as the target face (for 1 s, Figure 6).

The spatial arrangement of the surrounding faces was the same as in Experiment 3.1. Simultaneously for the same duration, the target face was presented at the center of the screen, superimposed under the fixation cross; therefore, the target face was presented onscreen within a group of 6 'friends'. Note that the duration of the target face in this experiment was longer than that of the adaptation experiments; however, such a setting allows the 'groups' of faces in each condition (i.e., adaptation or friends) to be presented for the same duration in each trial across experiments (1 second). Thus, the influence of the group of faces can be directly compared. A final screen with only the fixation was then be presented until participants pressed the appropriate keyboard response to indicate whether the target face was attractive or unattractive

('A' for attractive, 'S' for unattractive). This screen commenced with a 50 ms beep noise to alert participants to respond, with their response starting the next trial.

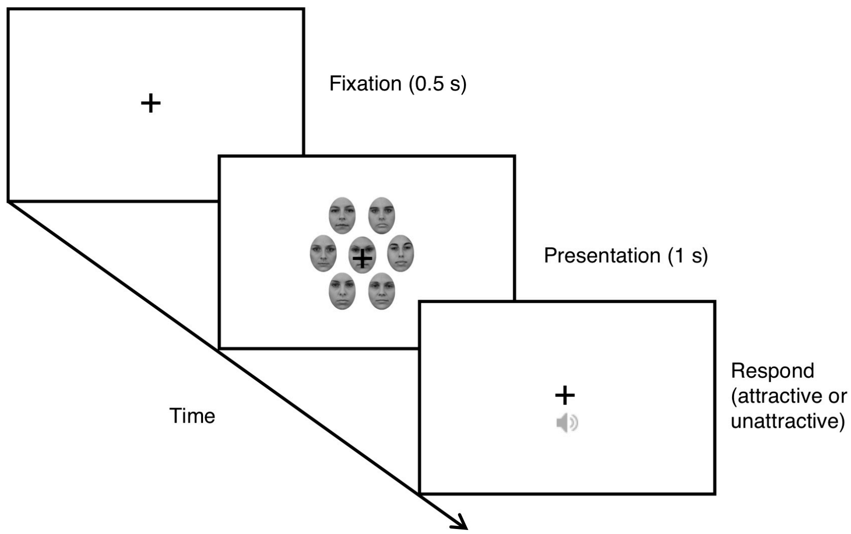

Figure 4.6. The sequence of one example trial (the demonstrated faces are AF01NES, AF06NES, AF08NES, AF17NES, AF20NES, AF26NES, and AF34NES from KDEF database). Subjects pressed the space bar to start a block. Then after 0.5 s, the target face, surrounded by the other six faces, appeared onscreen for 1 s. Then a beep sound prompted subjects to judge the attractiveness of the central face by pressing the 'A' button for attractive, or the 'S' button for unattractive. Experimental details can be found in the method section.

4.3.2. Results and Brief Discussion

The results from all participants averaged together are presented in Figure 4.7A. We plotted the fraction of 'attractive' responses as a function of the attractiveness units of the target faces. The friend effect can be interpreted from a leftward psychometric curve shift; whereby larger leftward shifts indicate a larger friend effect. The figure indicates that when the target faces were surrounded by either unattractive faces (solid blue line), mixed faces (cyan dashed line), or the attractive faces (magenta dotted line), all target faces were perceived as more attractive than on their own (baseline, black dash-dotted line). Moreover, it appeared that decreasing the average attractiveness of the surrounding group led to larger friend effects. The summary of the friend effect measured by PSE shift is illustrated in Figure 4.7B. The Shapiro-Wilk tests indicated that the data from the 'attractive surrounding' condition follows normal distribution ($p > .54$); however, the those from the other two conditions do not (*p*s < .05). Compared to the baseline PSE, the unattractive (M = 12.70%, $SEM = .027$; $Z = 3.77$, $p < .001$) , mixed (M = 9.21%, $SEM = .025$; $Z = 3.62$, $p < .001$), and attractive surrounding faces (M = 5.50%, $SEM = .021$; $t(19) = 2.61$, $p = .017$, Cohen's $d = 0.58$, 95% CI [0.01, 0.10]) all boosted the attractiveness of the centrally presented target faces. A Friedman test also indicated significant differences among all three friend conditions ($\chi^2(2) = 10.33$, $p = .006$). Further Wilcoxon signed-rank tests indicated that the unattractive condition produced the largest friend effect: larger than the attractive ($Z = 3.17$, $p = .002$) and larger than mixed friend ($Z = 2.09$ $p = .036$) conditions. Greater friend effects were found in the mixed over the attractive condition ($Z = 2.28$, $p = .023$).

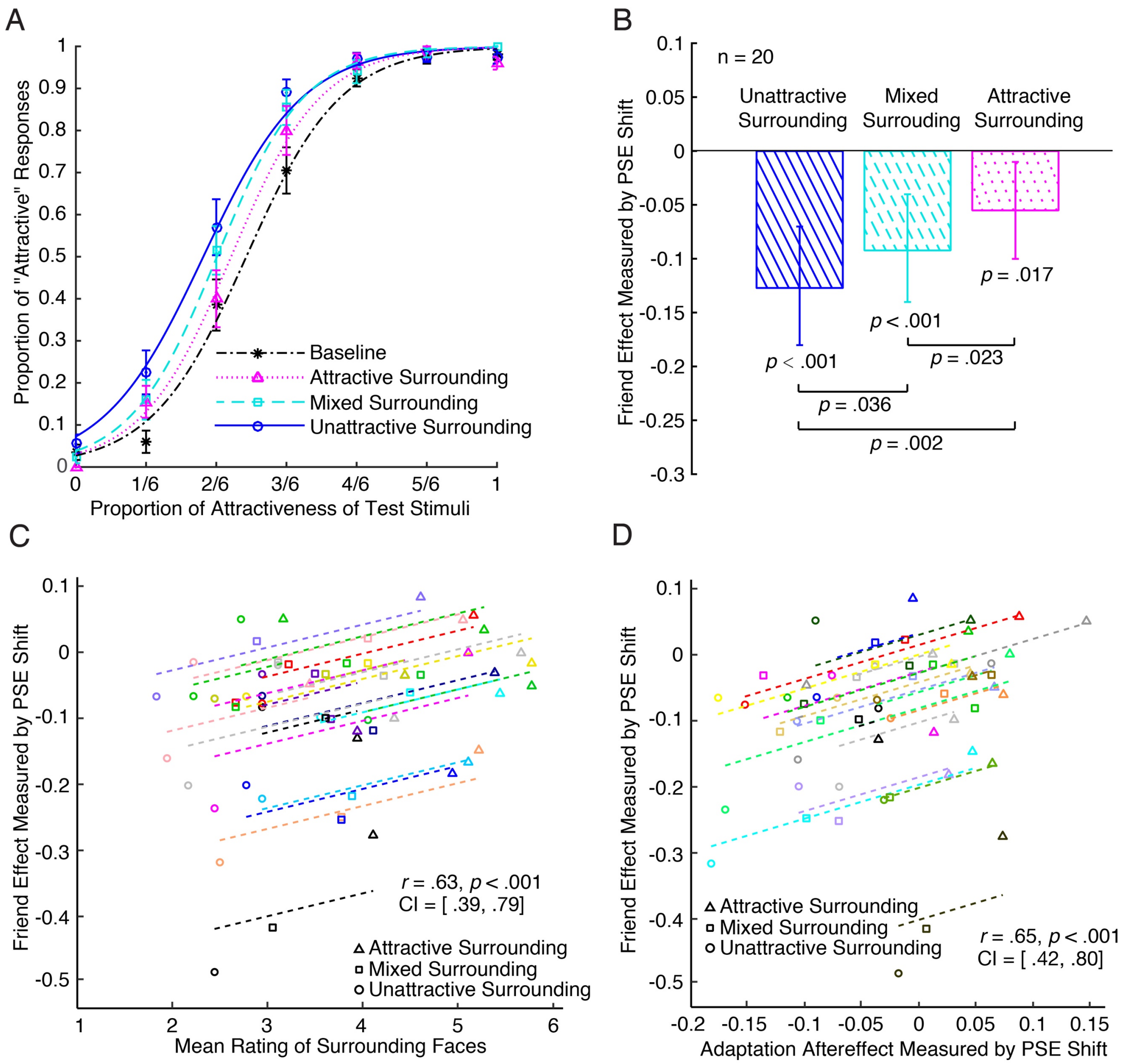

Figure 4.7. The effects of surrounding faces with different levels of group attractiveness (Experiment 3.3). (A) The psychometric functions of all participants averaged together. 'Baseline' represents the baseline condition without any surrounding faces (black star, black dash-dotted line). 'Attractive Surrounding' represents the attractive surrounding faces condition with six attractive faces (magenta triangle, magenta dotted line). Error bars represent the SEMs. 'Mixed Surrounding' represents the mixed surrounding faces condition with three attractive and three unattractive faces (cyan square, cyan dashed line). 'Unattractive Surrounding' represents the unattractive surrounding faces condition with six unattractive faces (blue circle, solid blue line). (B) Summary of all 20 subjects' results. For each condition, the Friend effect measured by PSE shift and the 95% confidence

intervals were plotted. The p-value above each bar was calculated against the baseline condition using the Wilcoxon rank sum tests, and one sample *t*-test. Note that, a more negative Friend effect measured by PSE shift indicates a larger friend effect (target faces were perceived as more attractive than on their own). (C) The friend effect as a function of the mean attractiveness rating of the surrounding faces. Magenta triangles represented the individual subjects' mean ratings of attractive surrounding faces. Cyan squares for the mixed surrounding, and blue circles for the unattractive surrounding. Taken together, the size of the friend effect and the mean attractiveness rating correlated significantly. (D) The magnitude of the attractiveness adaptation aftereffects in each condition (relative to the baseline) plotted as a function of the corresponding conditions' friend effects (calculated relative to the no friend baseline). Adaptation aftereffects and the friend effects were significantly correlated.

We specifically predicted that the friend effect might be influenced by ensemble perception by forming a new norm from the group's mean attractiveness ratings. To support this hypothesis (Figure 4.7C), we found a significant positive correlation between the mean attractiveness ratings for each surrounding group and the score of friend effect (r = .63, p < .001, 95% CI [0.39, 0.79]). In other words, target faces became more attractive as they were surrounded by less attractive friends. This indicates that the friend effect is influenced in a way that is consistent with ensemble adaptation.

To link the prior experience (Experiment 3.1) and contextual (Experiment 3.3) effects created via hypothesized ensemble coding, a further repeated measures correlation (Figure 4.7D) between the adaptation aftereffects in Experiment 3.1, and the friend effects from Experiment 3.3, was performed. The results showed that both were significantly associated with one another (r = .65, p < .001, 95% CI [0.42, 0.80]). This further confirms a link between ensemble adaptation and the friend effect, whereby both were occurring in a way that is consistent with the underlying ensemble representations of facial attractiveness (Figure 4.8). Both adaptation aftereffects and friend effects were significantly correlated with the mean attractiveness of the face group (r = .81, p < .001; and r = .63, p < .001). This suggests that the mean attractiveness of a face group is the common factor associated with both friend effects and attractiveness adaptation aftereffects.

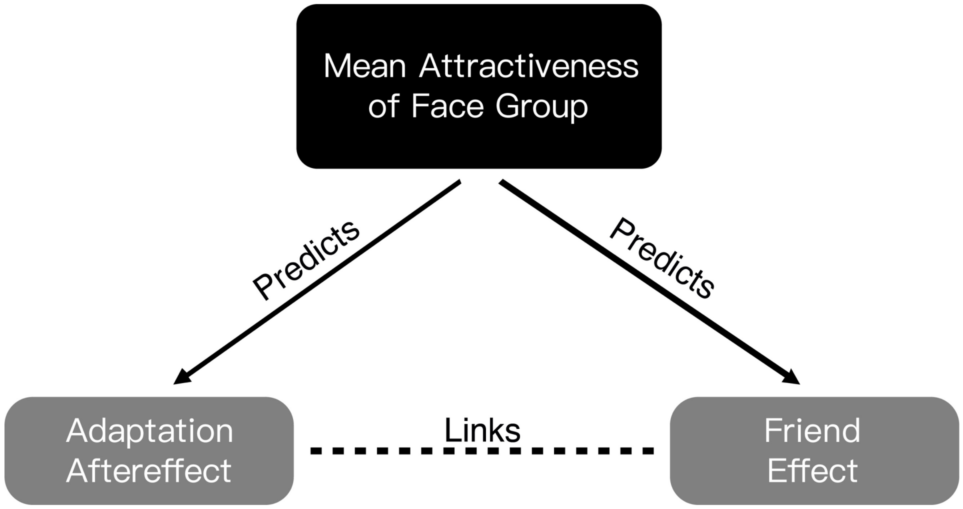

Figure 4.8. The generalized model from these experiments. The mean attractiveness of the face group (ensemble representation) could predict the adaptation aftereffect (Experiment 3.1) and the friend effect (Experiment 3.3). In the other words, the adaptation aftereffect and the friend effect were both associated with ensemble representations.

In summary, the results from Experiment 3.3 confirm that an individual's face is perceived as more attractive when it is presented with other faces than when presented alone. Our results therefore replicate that of prior work (Carragher et al., 2018; Walker & Vul, 2014; but see Ojiro et al., 2015).

In contrast to Walker and Vul's 'basking in reflected glory' theory of the friend effect, we find that the friend effect is negatively determined by the mean attractiveness of the surrounding faces: the more unattractive the friends are, the more attractive the target face becomes. The friend effect is therefore not a consequence of averaging the target face towards the group's mean, which should have made the faces more attractive in the attractive condition but seems to arise from a contrastive effect between the ensemble perception of the 'friends' and the 'target'. Moreover, if we consider the results from Experiment 3.1 where the 'Mixed Adapting' condition produced no adaptation aftereffects, the 'Mixed Surrounding' condition here could be reflective of a baseline friend effect. From this baseline, the attractive faces then diminish the effect and the unattractive faces boost it. Our results are therefore potentially consistent with the suggestion that the friend effect comprises of a social positive effect, where the 'Mixed' condition is the baseline of this effect, and a contrastive effect, which can then modulate the size of this social positive effect. This begs the question as to whether the friend effect is partly driven by the presence of surrounding faces, or requires variance between the faces too? To answer this question, we tested a new condition in Experiment 4, with identical faces surrounding the target face.

4.4. Experiment 3.4: Contrastive and variance, but not surrounding *per se*, determine the friend effect

In this experiment, we not only tried to replicate the findings in Experiment 3.3 with different groups of participants, but also tried to clarify the mechanisms behind the 'social positve' effect. We presented our participants with 'friends' that were identical to the targets (i.e., the target face was surrounded by six copies of itself); under such circumstances, the ensemble average of the 'friends' in the surrounding scene is the same as the target. We therefore anticipated that the contrastive effect would not be observed as there is no difference between the ensemble of the scene and the target, and so, only a social positive effect should occur when all faces in the scene are identical; e.g., the friend effect should be similar to the 'Mixed Surrounding' condition.

However, if the friend effect requires a contrast between the target and ensemble perception of the friends, then we might fail to find a friend effect when all faces are identical. Support for this possibility come from our adaptation results in Experiment 3.2, where a single face (Single1) in a scene produced no different aftereffects from a group of identical faces (Single6). If we consider that the friend effect arises in part due to similar neuronal activation as with adaptation, then we may not find a friend effect when all faces are identical. This is because we would expect the same neuron populations to be activated for a particular face or its identical copies. Either result would give us an important insight into what drives the friend effect. Moreover, we had groups of attractive, mixed and unattractive adapting faces similar to those used in Experiment 3.1 and 3.2, as well as test faces from Experiment 3.2, in order to replicate the friend effect we found in Experiment 3.3. Finally, we used different facial identities between adaptation and test faces in order to remove any possible confound of identity, much like how we changed the stimuli for Experiment 3.2 in order to counter the same issue in Experiment 3.1.

4.4.1. Methods

Subjects, Apparatus, Stimuli, and Procedure

Thirty new subjects participated in this experiment. We chose 30 as the sample size for several reasons. Firstly, using the effect size of Experiment 3.3 ($\eta_p^2 = .35$), power analysis (using G*Power 3.1, with α-value at .05, and power (1 – β) at .80) indicated that we need at least 13 participants. Considering the recent literature in attractiveness studies using similar paradigms (n_{mean} = 27.8 in Walker & Vul, 2014; n= 30 in Pegors et al., 2015), we believed 30 is sufficient for the experiment.

We used the same lab setting as in Experiment 3.3 and the updated stimuli as in Experiment 3.2, except for a small change to the paradigm. As mentioned before, apart from the replication of Experiment 3.3, we added the same surrounding condition to clarify the impact of social positive effect. Thus, in the new 'Same

Surrounding' condition, the target face was always surrounded by six copies of itself.

4.4.2. Results and Brief Discussion

The results from all participants averaged together are presented in Figure 4.9A. We plotted the fraction of 'attractive' responses as a function of the attractiveness units of the test faces. The friend effect can be interpreted from a leftward psychometric curve shift; whereby a larger leftward shift indicates a larger friend effect. The figure suggests that when the test faces were surrounded by either unattractive faces (solid blue line), mixed faces (cyan dashed line), or the attractive faces (magenta dotted line). All target faces were perceived as more attractive than on their own (baseline with no surrounding faces, black dash-dotted line). However, the 'Same Surrounding' condition has no obvious shift from the baseline condition, which means surrounded by the same face did not increase the attractiveness.

The summary of the friend effect measured by PSE shift is illustrated in Figure

4.9B. The Shaprio-Wilk tests indicate that all data follow normal distributions, *p*s > 0.33. Compared to the baseline PSE, the unattractive ($M = 12.8\%$, $SEM = .014$; $t(29) = 9.12$, $p < .001$, Cohen's $d = 1.67$, 95% CI [0.10, 0.16]) , mixed ($M = 8.59\%$, $SEM = .012$; $t(29) = 6.94$, $p < .001$, Cohen's $d = 1.12$, 95% CI [0.06, 0.11]), and attractive surrounding faces ($M = 4.5\%$, $SEM = .010$; $t(29) = 4.52$, $p < .001$, Cohen's $d = 0.59$, 95% CI [0.2, 0.7]) all boosted the attractiveness of the centrally presented test faces. However, the same face surrounding condition failed to invoke a significant PSE shift ($M = 1.4\%$, $SEM = .0093$; $t(29) = 1.46$, $p = .154$, Cohen's $d = .018$, 95% CI [-.005, 0.03]). A repeated measures analysis of variance (ANOVA) also indicated significant differences among all four conditions (Mauchly's test indicated that the assumption of sphericity was violated, $\chi^2(5) = 13.53$, p = .01; thus the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .74$); $F(2.23,64.66) = 35.12$, $p < .001$, $\eta_p^2 = .55$). Further pairwise comparisons indicated that the unattractive condition produced the largest friend effect: larger than the mixed (p = .002), attractive surrounding (p < .001) conditions, and the same surrounding condition (p < .001). Similarly, greater friend effects were found in the mixed over the attractive condition (p = .002) and over the same surrounding condition (p < .001). Finally, greater friend effects were found in the attractive over the same surrounding condition (p = .009).

We further validated whether the friend effect is influenced by ensemble perception by forming a new norm from the group's mean attractiveness ratings. Using repeated measures correlation analysis, we found a significant positive correlation (Figure 4.9C) between the mean attractiveness ratings for the three surrounding group and the score of friend effect (r = .54, p < .001, 95% CI [0.33, 0.70]).

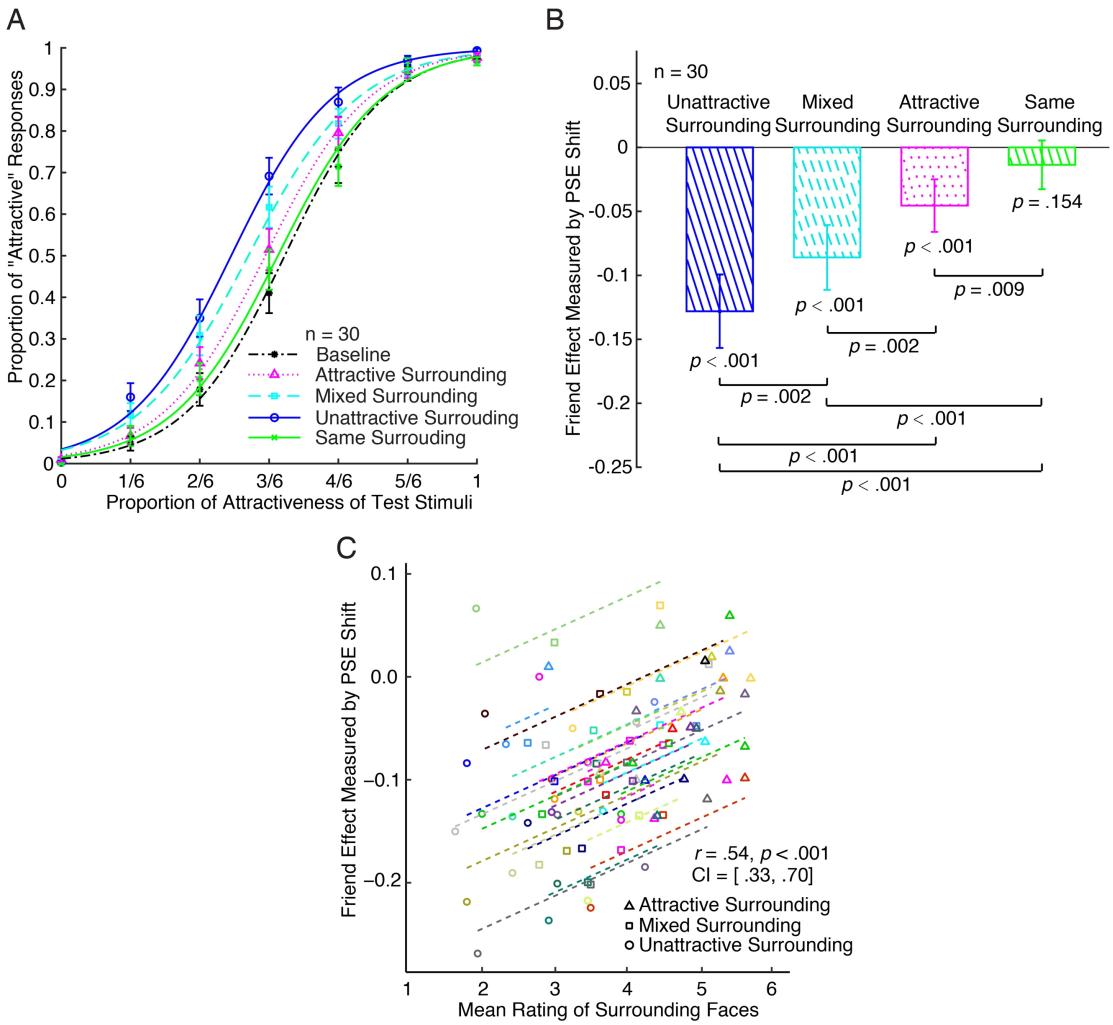

Figure 4.9. The effects of surrounding faces with different levels of group attractiveness (Experiment 3.4). (A) The psychometric functions of all participants averaged together. 'Baseline' represents the baseline condition without any surrounding faces (black star, black dash-dotted line). 'Attractive Surrounding' represents the attractive surrounding faces condition with six attractive faces (magenta triangle, magenta dotted line). Error bars represent SEMs. 'Mixed Surrounding' represents the mixed surrounding faces condition with three attractive and three unattractive faces (cyan square, cyan dashed line). 'Unattractive Surrounding' represents the unattractive surrounding faces condition with six unattractive faces (blue circle, solid blue line). 'Same Surrounding' represents the condition in which the target and 6 surrounding faces are identical (green X, solid green line). (B)

Summary of all 30 participants' results. For each condition, the Friend effect measured by PSE shift and the 95% confidence intervals were plotted. The p-value above each bar was calculated against the baseline condition using the pairwise comparison (Bonferroni corrected). Note that, a more negative Friend effect measured by PSE shift indicates a larger friend effect (target faces were perceived as more attractive than on their own). (C) The friend effect as a function of the mean attractiveness rating of the surrounding faces. Magenta triangles represented the individual subjects mean ratings of attractive surrounding faces. Cyan squares for the mixed surrounding, and blue circles for the unattractive surrounding. Taken together, the size of the friend effect and the mean attractiveness rating correlated significantly, as in Experiment 3.3.

Therefore, this study replicated the findings of Experiment 3.3. Moreover, it further clarifies that the 'friend effect' cannot be elicited by the mere presence of other, identical faces. Instead, it appears that there needs to be some variance between the faces in order for the friend effect to become engaged.

4.5. Discussion

In the first two experiments, the data showed that ensemble statistics of previously viewed groups could shape subsequent attractiveness judgments in the form of adaptation aftereffects. These judgments were correlated with the underlying mean attractiveness of the adapting group of faces, indicating that spatial ensemble perception was arising and producing attractiveness adaptation. Similarly, in Experiment 3.3 & 3.4, we tested whether the company we keep changes how others perceive our attractiveness. As expected, being surrounded by an increasingly unattractive group leads to you being more likely to be judged as attractive, causing a contrastive ‘bring out the beauty’ effect. The participants’ mean ratings of attractiveness of the surrounding faces were correlated with the size of their friend effects. Overall, it would seem that the brain can average the attractiveness of a group of faces together involuntarily, to form a new norm which target faces can be implicitly judged against. These findings may, therefore, indicate an evolutionary advantage in rapidly assessing a mate’s worth against past (adaptation) and present (friend effect) experiences. Overall, the adaptation and friend effect are two important face perception phenomena that are predicted in a fashion consistent with ensemble statistics.

Previous studies in facial attractiveness adaptation have tended to use configurally distorted faces as their adaptors (Rhodes et al., 2003). For instance, Anzures and Mondloch (Anzures & Mondloch, 2009) adapted children and adults to compressed and expanded faces to probe attractiveness perception. In this study, adaptors were natural faces without distortion, yet we still observed large aftereffects (similar to Webster et al., 2004; Webster & MacLeod, 2011). To our knowledge, the present study is the first that tests facial attractiveness adaptation through the use of natural, undistorted faces. Also, our aftereffects seem incompatible with a low-level retinotopic adaptation explanation (Afraz & Cavanagh, 2008; Leopold et al., 2001), as the adaptors and the test faces were presented in non-overlapping spatial locations in first two experiments. Such incongruence between adaptors and test face is typically thought to counteract low level retinotopic effects (Adams, Gray, Garner, & Graf, 2010; Leopold et al., 2001) and thus indicates the ensemble perception occurs at a higher level of face perception (Haxby & Gobbini, 2011; O'Doherty et al., 2003). Similarly, they support the suggestion that the perception of facial attractiveness is not entirely innate but can be shaped quite considerably by both context and experience (Ewing et al., 2010; Furl, 2016; Jones et al., 2007; Little et al., 2001; Little et al., 2011; Rhodes et al., 2003; Stormer & Alvarez, 2016).

In Experiment 3.2 we further confirmed that it is the ensemble coding of the crowd that drives adaptation aftereffects. Interestingly, the mere presence of multiple same faces in the crowd does not increase the adaptation aftereffect of a single face (Single1 vs Single6). Thus, these results clarify that the adaptation aftereffect derived from a crowd comes about through averaging, and not the summation of the individual faces. Moreover, by using different identities of adaptors and the target face (Experiment 3.2), we further clarified that the observed adaptation aftereffect can be only attributed to the facial attractiveness adaptation, rather than a consequence of facial identity adaptation.

Experiment 3.3 suggests that being in the presence of increasingly unattractive faces leads to greater friend effects, which is incongruent with the inferred prediction from Walker and Vul that attractive friends should make one more attractive (Walker & Vul, 2014). While the friend effect seems altered by ensemble perception, there is still a robust boost in the target face's attractiveness regardless of the surrounding faces' mean attractiveness. The findings of Experiment 3.3 are, however, still open to the interpretation that the friend effect is comprised of two components: a contrastive effect and a social positive effect. The results from Experiment 3.4 not only replicated the major findings from Experiment 3.3, but also clarified that the mere presence of faces (the 'Same Surrounding' condition) does not increase the attractiveness ratings of the central target face. Therefore, while the friend effect seems modulated by the contrast between the ensemble representation of the surrounding faces and the central target face, there needs to be variance between these faces (i.e., they cannot be identical) for the social positive component of the friend effect to become engaged.

Why should a face always be more attractive when viewed in a crowd?

Moreover, why does the attractiveness of the target face decrease when that of the surrounding faces increased? As mentioned earlier, we believe that the friend effect might have two components. The first component seems to be a social positive effect generated by the surrounding faces. The second is a contrastive effect between the target and the ensemble representation (mean attractiveness) of the surrounding faces. We therefore believe that the second component can be explained by ensemble neuronal habituation, similar to the ensemble adaptation aftereffects observed in Experiment 3.1 & 3.2. For example, prior work has shown that the specific neurons responsible for face perception in the inferior temporal cortex have large receptive fields and position invariance (Barraclough & Perrett, 2011; Gross et al., 1972; Desimone et al., 1984; Desimone, 1991; Tsao & Livingstone 2008). When the identical faces are presented at different locations ('Same Surrounding' condition), they may activate the same population of neurons as the single face in isolation (i.e., the 'Baseline' condition where there were no surrounding faces). This explains why there was no social positive effect in the 'Same Surrounding' condition in Experiment 3.4.

When the faces in a scene have variance by being different identities, we would expect each face to activate different populations of neurons from those of the baseline condition. We believe that when these additional face selective neurons are activated to detect multiple faces, it can allow the target to be appraised as more attractive because of this apparent popularity. While attending to the target face, and multiple faces have been detected in the scene, the brain can then engage a contrastive effect (Luck et al., 1997). This could explain why we always observe a friend effect, even when the friends are attractive. For example, in the 'Attractive Surrounding' condition, the social positive effect and the contrastive effect both occurred, but the social positive effect is always present due to the detection of variance in the faces preventing the contrastive effect from eradicating it entirely or reversing it. Future neuroimaging work will, however, be required to clarify the neural mechanisms of the social positive effect.

It should be noted that we are not claiming ensemble coding cannot occur when there is no variance in the scene. For example, Luo and Zhou (2018) recently showed variability is not required for ensemble perception of facial attractiveness to arise. Instead, for the friend effect to occur, variance is required. We have demonstrated this through the 'missing' friend effect in the identical face (the 'Same Surrounding' condition) in Experiment 3.4. Similarly, there appeared to be nothing special with respect to adapting to an ensemble of identical faces when compared to the single face in Experiment 3.2. Based upon our current experiments, however, we are unable to answer whether or not the target face is included in the ensemble representation.

The observed friend effect seems similar to the center-surround inhibition in low-level vision: the appearance of other stimuli around the target 'suppress' the perception of the target stimuli. However, we believe the friend effect is not likely to be a process which is similar to center-surround effects in low-level visual dimensions. The center-surround inhibition describes the fact that the response of a neuron (e.g. in the medial temporal visual area) to a favored stimulus in the classical receptive field would be inhibited by the simultaneous appearance of other stimuli outside of the classical receptive field (e.g. Allman, Miezin, & McGuinness, 1985). However, in our Experiment 3.3 and 3.4, the spatial arrangement of the stimuli ensured that the faces were within the classical receptive fields of the face processing neurons: face perception related neurons in the inferior temporal cortex have large receptive fields and position invariance (Barraclough & Perrett, 2011; Gross et al., 1972; Desimone et al., 1984; Desimone, 1991; Tsao & Livingstone 2008). Moreover, if the friend effect is a face-level center-surround inhibition, then the 'same surrounding' condition should generate an attractive boost. However, the insignificant friend effect from the 'same surrounding' condition further suggests that the mere appearance of faces around the target face does not alter the friend effect, and thus the observed friend effect is not face-level center-surround inhibition. Future research may consider using other direct measurements to further clarify the similarity and difference between the friend effect and the center-surround inhibition.

In Experiment 3.2 the ‘Single1’ condition was quite different from the other three adaptation conditions: the other three conditions had six faces in the adaptor, while the ‘Single1’ condition had one intact face and five scrambled facial images. Such a design offered us a chance to directly compare the adaptation aftereffect of one face (‘Single1’) against that of a group of the same faces (‘Single6’) and clarify whether the observed adaptation aftereffects in Experiment 3.1 was an averaging (ensemble) or a summation (a cumulative effect caused by multiple faces) process. The observed similar adaptation aftereffects supported the ensemble averaging notion. We have to admit, to a certain extent, that the participants may have casted their attention in different manners in ‘Single1’ and ‘Single6’ conditions. For instance, the single intact face in the ‘Single1’ condition should attract more attention than any of its copies in the ‘Single6’ condition: faces attract attention in an automatic fashion, and our visual system is biased to faces over other visual stimuli (Palmer & Rhodes, 2007). Therefore, it is reasonable to say that the participants tend to pay more attention to the single intact face over the five scramble images. However, the design itself was to clarify the mechanisms behind the observed aftereffects. So, to examine the so-called 'summation' hypothesis, we have to check whether the aftereffect of the 'Single6' condition is a summation of adaptation aftereffects to each individual face (six times of the aftereffect of 'Singe1').

It is remarkable that ensemble perception occurs in adaptation and the friend effect experiments even though participants were never instructed to look directly at the group of surrounding faces, and in the case of Experiments 3.3 and 3.4, only paid attention to the single target face in the center of the face crowd. This suggests that ensemble perception can occur regardless of directed attention and supports the claim that such perception is an involuntary process. Similarly, the fact that our participants never look directly at the faces would seem to indicate that they were not simply ‘picking’ a single face out of the group to base their judgments upon. This suggestion is further supported by the correlations between the mean attractiveness ratings of the underlying groups and the friend/adaptation effects. Our current study supports the suggestion that ensemble perception occurs for facial attractiveness adaptation, in addition to emotion adaptation (Ying & Xu, 2017). These findings fit with the view that ensemble perception likely arose through an evolutionary advantage at being able to judge the attractiveness of any particular face against recent and current experiences. However, such high-level ensemble coding is likely to have arisen from a precursor system that initially processes lower level information such as contours and textures in the environment. One could imagine that as man became a more social creature, then this process was developed for complex, higher level social information such as attractiveness. In any case, this perspective suggests that other aspects of facial social traits (Oosterhof & Todorov, 2008), such as trustworthiness or dominance, might also be susceptible to ensemble encoding.

In summary, this study with four experiments has provided evidence that supports the ubiquitousness of ensemble coding in shaping face perception. The converging data has further confirmed the robustness and replicability of the friend effect, whereby simply being in the company of friends, at least different looking from the target, makes an individual look more attractive. Similarly, if the viewer has prior exposure to unattractive groups of faces, then a face will again seem more attractive when subsequently viewed by comparison. If individuals therefore want to maximize their mating competitiveness by seeming more desirable, they should surround themselves with unattractive friends or appear after them.

Chapter V: Study 4: Attention Modulates the Ensemble Coding of Facial Expressions

What is the relationship between attention and ensemble statistics? Recent evidence in ensemble coding of size suggested that the attended items contributed more to the averaging (Chong & Treisman, 2005; de Fockert & Marchant, 2008; Li & Yeh, 2017). Moreover, some new evidence showed that the perceptual averaging requires attention: when attention was biased to the secondary task, subjects were unable to average the emotion between faces (Elias et al., 2017). Thus, it is possible that attention would modulate high-level ensemble statistics.

However, there are reasons to be skeptical that attention could interfere with the ensemble statistics of facial expressions. Facial expressions are processed through both cortical and subcortical pathways (Haxby & Gobbini, 2011; Haxby et al., 2000, 2002; Johnson, 2005; Kleinhans et al., 2011; Lovejoy & Krauzlis, 2010; Nguyen et al., 2016): cortical regions like superior temporal sulcus (STS), insular, orbitofrontal cortex (OFC); subcortical regions like amygdala, pulvinar, superior colliculus (SC). A substantial body of evidence from brain imaging studies suggests that face processing in the amygdala is hardly affected by attention modulation, while cortical face processing (e.g. STS) is gated by spatial attention (Holmes et al., 2003; Vuilleumier et al., 2001). Consequently, attention might not necessarily impact the ensemble coding of facial expressions. If the ensemble coding of facial expressions occurs in the amygdala but not the other attention sensitive face processing regions, then attention would not impact the ensemble coding, and *vice versa*. Therefore, examining the relationship between attention and the ensemble statistics of facial expression might not only clarify the computation of ensemble coding but also hint at the neural mechanism of the ensemble coding.

In brief, this study aims to clarify that whether and how attention modulates the ensemble coding of facial expressions. The first experiment tested whether the exogenous cues affect the reported mean emotion of a crowd of faces. Then, the second experiment further examined the role of attention using endogenous cues and adaptation paradigm.

5.1. Experiment 4.1: Attention modulates the reported mean emotion of a group of faces

The current experiment aims to study how attention to a single face affects the perceived mean emotion of the crowd, by using exogenous cues. Subjects were cued either to the happiest face, the saddest face, or simply at the central fixation (not to any faces).

If the perceived mean emotions are similar among different cueing conditions, then this would suggest that ensemble coding is merely an averaging of the crowd, regardless of attention. Thus, the ensemble statistics of faces might occur in amygdala. Alternatively, if the perceived mean emotions are significantly different among cueing conditions, then it would suggest that ensemble coding is subject to the loci of spatial attention. Thus, the ensemble statistics of faces might occur in the attention sensitive face processing regions (e.g. STS, SC, pulvinar).

5.1.1. Methods

Subjects

Ten subjects (5 females, mean age 26.1), with normal or corrected-to-normal vision participated this study. This number of subjects is based on previous studies on attention and ensemble coding of sizes (de Fockert & Marchant, 2008; Li & Yeh, 2017).

Stimuli

This experiment selected facial identity AM14 from Karolinska Directed Emotional Faces (Lundqvist et al., 1998) database as the testing stimuli. Pictures with Happy, Neutral, and Sad expressions (AM14HAS, AM14SAS, and AM14NES) from this chosen identity were selected. Then we used Webmorph software (DeBruine & Tiddeman, 2017), and Matlab (Mathworks, Natick, MA) to morph and further manipulate these faces. All of the faces were grey scaled and cropped by an oval-shaped mask with only the central region of each face remaining visible. The luminance and contrast of the faces were matched by SHINE toolbox (Willenbockel et al, 2010.). All test stimuli were 2.40° × 3.02° in size.

We created one continuum of emotional faces with emotions from Sad to Neutral to Happy respectively to create the testing stimuli (Fig 5.1). The happiest faces showed 100% of happiness, the neutral faces showed 50% of happiness, and the four saddest faces showed 0% of happiness.

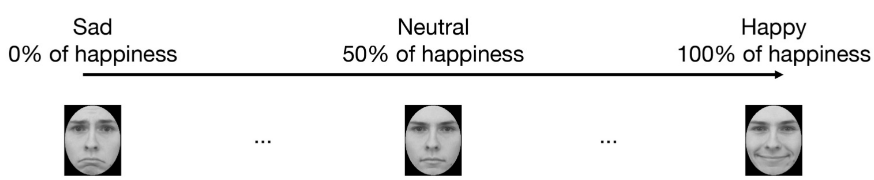

Figure 5.1. Stimuli used in experiment 1. The continuum of emotional faces from sad (0% of happiness), to neutral (50% of happiness), to happy (100% of happiness).

The target stimuli are four faces from the continuum. Their mean emotion was either 45%, 50%, or 55% of happiness. The four faces were +45%, +15%, -15%, and -45% of happiness compared to the mean. Therefore, there was one happy face, one sad face, and two mediocrely neutral faces. The test faces are seven faces from the continuum. They have 0%, 20%, 35%, 50%, 65%, 80%, and 100% of happiness. We presented all of the test faces on the screen simultaneously (Figure 5.2) on a black background. Each face corresponds to one button.

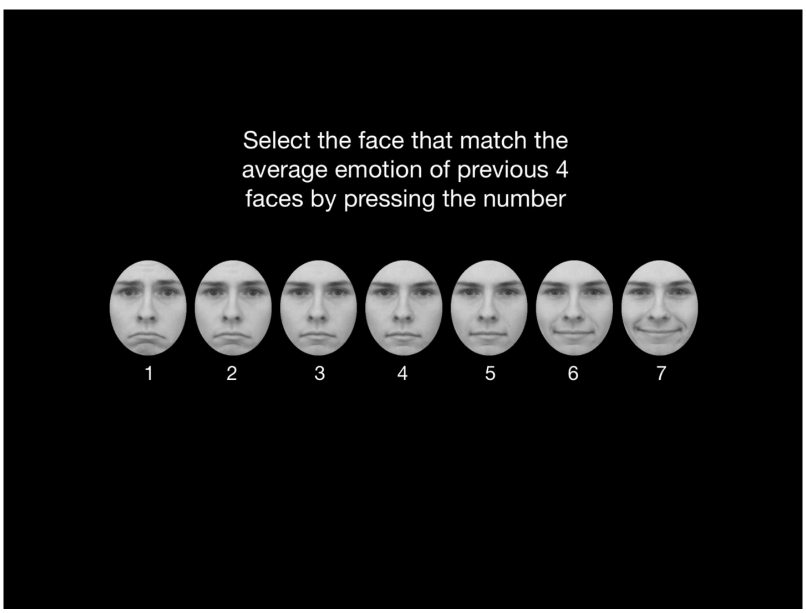

Figure 5.2. The testing phase of experiment 1. Subjects see this screen for unlimited amount of time. They were asked to select the face, by pressing the correspondent button, with the mean emotion of the target faces.

Procedure

Before the actual experiment, subjects went through a practice section with 4-8 trials to become familiar with the procedure of the experiment. Then they commenced the experiment.

The scheme of one trial is illustrated in Fig 5.3. Each trial commenced with a fixation (1000 ms). Subjects were forced to concentrate on it until it disappears. Then a visual cue appeared on the screen for 188 ms. It might appear at the location of the happiest face (which presented later) for 1/3 of the time; or at the location of the saddest face (which presented later) for 1/3 of the time; or at the location of the fixation cross for 1/3 of the time. After a 94 ms interval, four target faces appeared on the screen for 1000 ms. At last, subjects were asked to select one face, by pressing the correspondent button, that represented the mean emotion of the target faces as soon as possible.

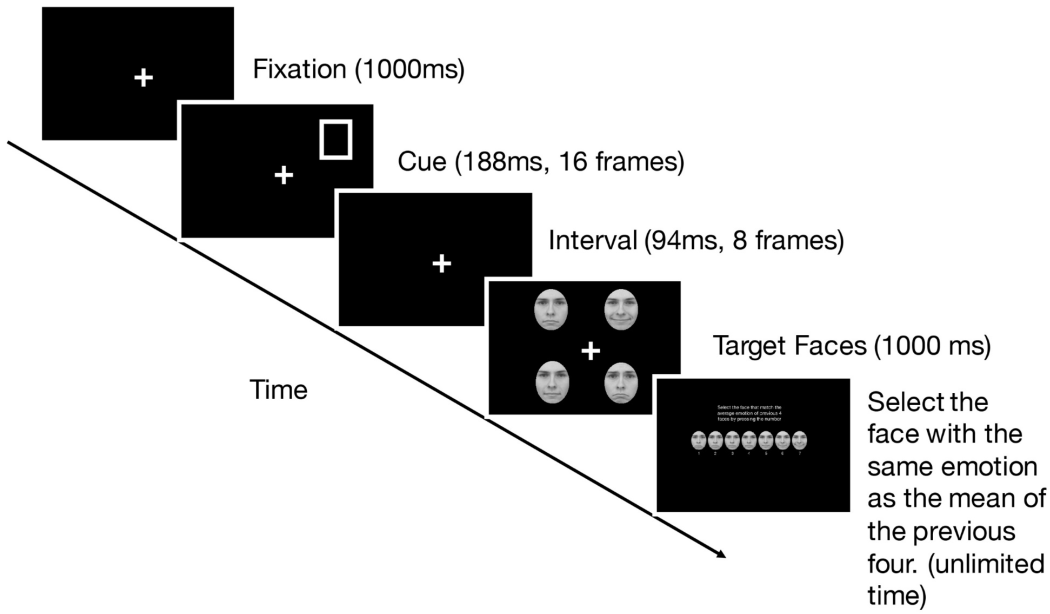

Figure 5.3. The trial sequence of one example trial in experiment 4.1. In each trial, the experiment initiates with a 1000 ms fixation stage. Then a cue appears on screen for 188 ms. In this example, the cue aims to guide subjects' attention to the happiest face.

Analysis

We plotted subjects' response distributions for each cueing condition individually. The x-axis represents the emotion value that the subject selected as the mean emotion of the target crowd (corrected from the raw data); and the y-axis represents the frequency of that selection. We then fit these responses with a Gaussian distribution. The x value, whereby the peak of the distribution occurs, is used to estimate the subject's perceived mean emotion in that condition. Afterward, we measured the cueing effect by calculating the difference between the perceived mean emotion of two cueing conditions against that of the baseline condition (the neutral cue condition).

5.1.2. Results

To clarify the impact of attention on the ensemble coding of the facial expressions, we compared the shift of perceived mean emotion of all subjects from baseline (cued to fixation condition). The Shapiro-Wilk tests suggested that the data from 'Cueing to Happy' condition does not follow normal distribution (p = .03), while that from 'Cueing to Sad' condition follows normal distribution (p = .17). Results of Wilcoxon rank sum test and result of t-test illustrated that cueing to the happiest face (M = 7.67%, SEM = 2.75%; Z = 64, p = .003) and cueing to the saddest face (M = -8.91%, SEM = 3.29%; $t(9)$ = - 2.71, p = .022, Cohen's d = 0.82) significantly increased the weight of the attended face in the reported mean emotion of that crowd, compared to the baseline condition (Fig 5.4.). Also, we found a significant difference between these two conditions (Z = 65, p = .002).

Figure 5.4. Summary of all subjects' results from Experiment 4.1. For each condition, the effect of cueing was measured as a reported mean emotion shift relative to baseline (cued to the fixation cross; error bars = SEM). The p values for each condition in the figure were calculated by Wilcoxon rank sum test and one sample t-test, comparing each cueing condition with the baseline or between cueing conditions.

5.1.3 Brief Discussion

The effect of the exogenous cues on the ensemble statistics of facial expressions was measured in this experiment. As hypothesized, the exogenous cues influence the reported mean emotion of the crowd by elevating the weight of the cued face. When subjects were cued to the happiest face, their perceived mean emotion of the crowd was happier; while they found the crowd sadder when cued to the saddest face. The findings here are consistent with previous research which showed that attention modulates the ensemble statistics of size (Chong & Treisman, 2003; Chong & Treisman, 2005; de Fockert & Marchant, 2008; Li & Yeh, 2017).

5.2. Experiment 4.2: Attention modulates the adaptation aftereffect of a group of faces

Results from Experiment 4.1 indicate that the explicit and voluntary ensemble perception of facial expressions is subject to the locus of spatial attention. Is this effect confined to the specific paradigm (namely the exogenous cueing and explicit report)? Could top-down attention impact the implicit averaging of facial expressions?

In this experiment, we use endogenous cues and an adaptation paradigm to address these questions (Corbett, Wurnitsch, Whitney, Schwartz, & Whitney, 2012; Ying & Xu, 2017). By doing so, the ensemble representation of the faces could be measured by the adaptation aftereffect. If attention cannot alter the implicit ensemble statistics of facial expressions, we should observe the indistinguishable adaptation aftereffects from two cueing conditions and the no-cue condition. Alternatively, if spatial attention modulates the adaptation aftereffect, we should observe distinctive aftereffects biased by the emotion of the cued faces.

Since face processing is hierarchical, adaptation to low-level features is possible to generate an aftereffect at higher-level face perception (Xu, Dayan, Lipkin, & Qian, 2008). Therefore, in this experiment, we presented adaptors and the testing stimuli at different sizes and arranged them at spatially offset locations to minimize the effect of the lower-level features in the adaptation aftereffects (Afraz & Cavanagh, 2008; Webster, 2015; Zhao & Chubb, 2001). Thus, the observed aftereffects in this experiment can only be attributed to high-level face adaptation aftereffects.

5.2.1. Methods

Subjects, Apparatus, Stimuli

Eight new subjects, with normal or corrected-to-normal vision, participate in this study. The apparatus and stimuli are adapted from Experiment 4.1. The most significant difference here is the happy and angry expressions of the facial identity AM 14 from KDEF database was selected. We also applied the same stimuli manipulation from the previous study on the raw pictures.

The adaptors here were the happy and angry faces from the selected identity (AM 14), which were further processed in the same way as the stimuli in Experiment

1. We then created a morph continuum from happy to neutral to angry expressions for the testing faces (Debruine & Tiddeman, 2017). We chose angry instead of sad in the current experiment to further validate that the role of attention is not limited to a certain expression pair. We quantified the emotion of each morph by the proportion of happiness in the face. In the continuum, the happiest face carries 100% of happiness, the neutral face carries 50% of happiness, and the angriest face carries 0% of happiness. The faces with 70, 60, 55, 50, 45, 40, 30% of happiness were selected as testing stimuli. Noticeably, the testing faces were set at the size of 75% of the adaptors ($2.40° \times 3.02°$) to minimize the low-level aftereffects (Burton, Jeffery, Calder, & Rhodes, 2015; Zhao & Chubb, 2001).

Procedure

The general procedure of this experiment is similar to previous studies testing facial adaptation aftereffects (Figure 5.5). Each subject completes three adaptation conditions and the baseline condition, which were in randomized orders. In all adaptation conditions, the adaptors are two happy faces and two angry faces. The adaptors surround the central cross in a square fashion (at 3° eccentricity). The happy faces and angry faces are always presented at one side of the screen with counterbalanced orders, which means in half of each adaptation condition, the happy faces appear at the left part of the screen while the angry faces appear on the opposite half. The three adaptation conditions contain the same adaptors but with one difference which is the type of endogenous cue. In the no-cue adaptation condition, there is no cue; in the happy and angry cue conditions, the endogenous cue always points to the location of the corresponding expression. The cue is a white arrow at 1° above the central fixation cross.

During the whole experiment, subjects were asked to fixate only at the fixation cross. Each trial starts with a 506 ms fixation. Then the cue appears for 1000 ms (for the no-cue condition, just a fixation cross). Then the adaptor appears with the cue for 200 ms. After a 94 ms interval, one testing stimulus appears for 47 ms. Then subjects are allowed to judge the emotion of that face by pressing ‘A’ (happy) or ‘S’ (angry) on the keyboard. In the baseline condition, there is no adaptation phase but only the target stimuli and the judgment.

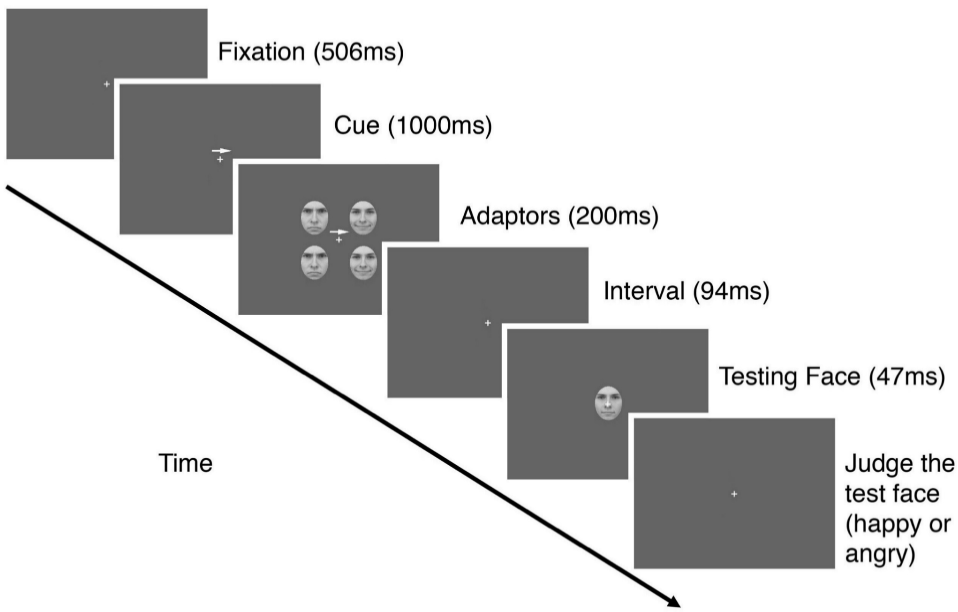

Figure 5.5. The trial sequence of one example trial in Experiment 4.2. In each trial, the experiment initiates with a 506 ms fixation stage. Then an endogenous cue appears on screen for 1000 ms. In this example, it is the happy cueing condition, and the cue points to the happy faces. After that, the adaptors appear on the screen for 200 ms. After a 94 ms interval, the test faces appear for 47 ms. Subjects then judge the emotion of the target face by pressing the corresponding button.

Analysis

Adapted from that of Study 1. Note that, in this experiment, the ‘No Cue’ condition is not the baseline for the adaptation aftereffect. It serves as a control to make sure whether attention could affect the inner representation of the emotional facial expressions. The actual baseline condition is the one with no adaptors, but just the testing faces.

5.2.2. Result

We compared the mean PSEs relative to the baseline of all subjects to measure the adaptation aftereffect (Figure 5.6). Results from Wilcoxon rank sum test showed that the no-cue condition (Shapiro-Wilk test suggests the aftereffects were not normally distributed, p = .024) failed to induce a significant attractive aftereffect (Z = 13.00, p = .547). However, results from t-tests (Shapiro-Wilk tests suggest the two aftereffects were normally distributed, ps > .40) suggested adaptation aftereffects in cue-to-happy condition (M = -2.45%, SD = 2.62%; $t(7)$ = - 2.64, p = .03, Cohen's d = 0.93) as well as cue-to-anger condition (M = 2.50%, SD = 2.44%; $t(7)$ = .2.87, p

= .02, Cohen's d = 1.01) generated significant aftereffects, but pointing in opposite directions.

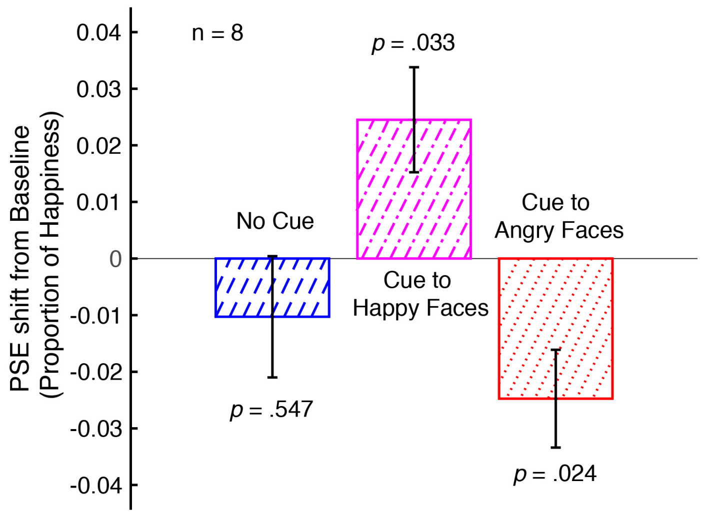

Figure 5.6. Summary of all subjects' results of Experiment 4.2. For each condition, the effect of cueing was reflected by the adaptation aftereffect (PSE shift from baseline condition). The blue bar with a dashed hatch line represents the no-cue condition; the magenta bar with a dash-dotted hatch line represents the cue-to-happy condition; the red bar with a dotted hatch line represents the cue-to-sad condition. The p values for all condition was calculated by Wilcoxon rank sum test paired *t*-tests, comparing the PSE of each cueing condition with that of the baseline. All error bars indicate the SEM.

5.2.3. Brief Discussion

This experiment employed endogenous cues and an adaptation paradigm. The results suggested that the locus of attention modulates the implicit ensemble statistics of facial expressions. The insignificant aftereffect observed in the no-cue condition here actually confirms the existence of ensemble statistics: the visual system does average the facial expressions together, thus the ensemble representation is precisely the average between happy and angry, which could not induce an adaptation aftereffect (Burton et al., 2015). Therefore, the ensemble representation was involuntarily calculated during the adaptation stage. Noticeably, the short presentation duration of the face crowd (200 ms in the current experiment) was sufficient for implicit ensemble statistics, which converges with other studies (Whitney & Leib, 2018).

Furthermore, the short duration of adaptors as well as the spatial offset between the adaptors and the target face together minimized the contribution of the low-level features in the observed adaptation aftereffects (Adams, Gray, Garner, & Graf, 2010; Afraz & Cavanagh, 2008; Bi, Su, Chen, & Fang, 2009). Consequently, the observed aftereffect could not be explained by some low-level features of the faces (only), but by the high-level adaptation to the ensemble representation of facial expressions.

5.3. Discussion

This study, with two experiments, employed different paradigms and scrutinized the role of attention in the ensemble statistics of facial expressions. Converging evidence showed that the emotion of the face(s) within the locus of attention heavily influences the ensemble representation of the crowd. When the subjects were required to report the mean emotion of the crowd, the emotion of the attended face biased the perceived mean (Experiment 4.1). Similarly, the location of the endogenous cue impacted the implicit ensemble averaging of the crowd, which was reflected in adaptation aftereffect (Experiment 4.2). Therefore, the ensemble representation is not calculating a simple arithmetic mean of the stimuli. It is more likely to be weighted averaging.

As the role of attention has been revealed in both experiments, it is highly likely that the ensemble statistics of facial expression are conducted in the attention sensitive regions for emotion perception (Holmes et al., 2003; Vuilleumier et al., 2001). Thus, the core facial processing regions (e.g. STS) could be the origins of the ensemble statistics of facial expressions (Haxby & Gobbini, 2011; Haxby et al., 2000). The SC and the pulvinar are also candidate regions (Kleinhans et al., 2011; Lovejoy & Krauzlis, 2010; Nguyen et al., 2016). Although it has been shown that the ensemble statistics abilities of high- and low-level features could not predict each other (Haberman et al., 2015), current findings together with a growing body of evidence suggests that the ensemble statistics of high- and low-level objects are both subject to attention modulation. Considering the fact that face perception is hierarchical (Xu et al., 2008), future research in neuroscience may further examine the possible neural mechanisms between high- and low-level ensemble statistics.

In Experiment 4.1, subjects were cued to one of the faces which has the most extreme emotion in the crowd. This method is similar to that of a previous report (de Fockert & Marchant, 2008) whereby the authors instructed subjects to attend to either the smallest or the largest circle in the set. Their study as well as the current study both find that attention modulates the mean representation by altering the weight of the attended item in averaging. However, researchers in that study modulate the attention differently in their two experiments: they specify the size of the circle to attend to in their first experiment and highlight one circle by changing the luminance in the second experiment. This made certain that they explicitly measured the ensemble statistics, which is also what this study ensured. The similarities in the experimental design and the observed findings suggest that the modulation by attention of ensemble statistics is ubiquitous and somewhat fundamental in ensemble statistics.

Typically, studies measure the attention's modulation on ensemble representation explicitly. For example, Li and Yeh (2017) asked their subjects to adjust a circle to match the mean size of the earlier shown circles. Conversely, Experiment 4.2 used the adaptation paradigm to measure the involuntary ensemble statistics. The adaptation paradigm is a powerful method to study the representation of stimuli in our visual system (Webster, 2015; Ying & Xu, 2017). The subtle change in perception which is undetectable via explicit measurement could be reflected by an adaptation paradigm (Luo et al., 2015). The subjects from Experiment 4.2 were not instructed to process the crowd, but the implicit ensemble representation still generated significant adaptation aftereffects on the subsequently viewed testing faces (Ying & Xu, 2017). Moreover, the findings of Experiment 4.2 further indicate that the observed effect in Experiment 4.1 is not confined to the specific experimental controls. To our knowledge, this is the first direct evidence suggesting that attention affects the involuntary ensemble statistics of faces.

In this study, the variances between emotions are large. Previous researchers studied the ensemble statistics with relatively small variances between objects. For instance, Haberman and colleagues (2009) tested the temporal ensemble statistics of facial expressions with four faces. The emotion difference between faces in their first experiment equals 6% of the emotional units in the current study. Unlike them, the emotion variance in Experiment 4.2 is much bigger (at 100%). However, it is reasonable to believe such a huge variance did not obscure the ensemble coding. Elias and colleagues (2017) found that the visual system is capable of averaging the emotion between fully happy (or angry) and neutral faces (50% in emotional units). Besides, another study showed that a stream consisting of half happy and half sad faces could be averaged together (Ying & Xu, 2017). They also suggested that the variance of emotion does not alter the perceived mean emotion.

The insignificant aftereffect in the no-cue condition from Experiment 4.2 confirms the existence of ensemble statistics: the visual system does average the facial expressions together, so the ensemble representation is precisely the average between happy and angry, which could not induce adaptation aftereffect (Burton et al., 2015). Thus, the ensemble representation is involuntarily calculated during the adaptation stage. Furthermore, the short duration of the presentation of adaptors as well as the spatial offset between the adaptors and the target face minimized the effect of low-level adaptation. Consequently, the observed aftereffect could only be explained by the high-level adaptation to the ensemble representation of facial expressions, and not to some low-level features of the faces.

To summarize, converging evidence from the two experiments showed that ensemble representation of facial expressions is sensitive to the emotion of the cued face(s). The modulation of attention occurs for both explicit and implicit ensemble statistics of facial expressions, by an alteration of the weight of the cued faces in the perceptual averaging. The findings here indicate that weighted averaging is an important characteristic of ensemble statistics of faces.

Chapter VI: Conclusion

6.1. General Discussion

In this thesis, four studies with thirteen experiments systematically examined the mechanisms of ensemble statistics of faces. These four studies: (1) suggested distinctive mechanisms of temporal and spatial ensemble statistics; (2) showed the prevalence of ensemble statistics in face perception; (3) linked ensemble statistics with important phenomena in face perception; and (4) showed the relationship between attention and the ensemble statistics of faces.

The results from Study 1 suggested that the emotions of the sequentially presented faces are involuntarily perceived via ensemble statistics. Further controls in emotion variance, temporal frequency, and facial identity suggested the robustness of the ensemble coding. As long as the mean emotions are the same, regardless of variance, the adaptation aftereffects are the same.

Since the average of the unattractive faces is (moderately) more attractive than any of its components, Study 2 examined the computational mechanisms of ensemble statistics, the fundamental question of ensemble statistics which has not yet been fully clarified. The results suggested that the temporal and spatial ensemble statistics of faces stem from different mechanisms. The temporal ensemble coding extracts the low-level 'computational' information from the faces, while the spatial ensemble statistics is averaging the high-level gist.

After confirming the involuntariness of ensemble statistics in facial expression, Study 3 further examined the ubiquitousness of ensemble coding in face perception. Results from four experiments together showed that two important phenomena related to facial attractiveness perception, namely adaptation and the cheerleader/friend effect, are both shaped by and linked with ensemble perception. The human visual system involuntarily updates the facial attractiveness perception by utilizing the ensemble statistics over the previous and current experience of other's faces.

The thesis then further studied the mechanisms behind ensemble statistics by examining its relationship with attention. The results from two distinctive paradigms here showed that attention plays an important role in perceptual averaging. The attended object has a bigger weight in the averaging. Thus, the ensemble coding of faces is a weighted average. These studies together also indicated the ubiquity and robustness of ensemble statistics in face perception.

Taken together, findings from the four studies expand current understanding of visual perception and face processing. The following sections specifically discuss the important findings from this thesis and compare the studies here with relevant works.

6.1.1. Mechanism of Ensemble Statistics of Faces

Our visual system is capable of extracting a summary representation of the facial characteristics provided by a group of faces (Haberman & Whitney, 2007, 2012; Sweeny & Whitney, 2014; Whitney & Leib, 2017; Ying & Xu, 2017). We do not fully understand the detailed mechanisms of it. This renders three questions: (1) what is 'averaged' during the ensemble statistics of faces? (2) what could affect the ensemble coding? (3) what are the possible mechanisms for ensemble statistics of faces? This thesis addressed all these questions.

Most studies in the literature describe the ensemble statistics as 'averaging the gist', implying that a general representation of the perception of the group is extracted (Haberman & Whitney, 2012; Whitney & Leib, 2017). However, it is possible that only the low-level features of the stimuli were extracted. Previous studies implied that it is impossible to differentiate between these two hypotheses using facial expression and identity. For facial attractiveness, however, the low-level 'computational' averaging theory produces a conflicting prediction as to what will be perceived when compared to the group's high-level gist average (DeBruine et al., 2007). Therefore, studying the ensemble coding of facial attractiveness benefits not only the understanding of facial attractiveness but also the understanding of ensemble statistics itself.

By examining the ensemble perception of facial attractiveness, Experiment 2.1 & 2.2 showed asymmetrical adaptation aftereffects between the attractive and unattractive face streams. The results suggested that contrary to the pervasive gist hypothesis, the visual system actually extracts the low-level 'computational' information when each face in the group is presented individually across time: the ensemble representation of temporally presented unattractive faces is more attractive than any component. However, using the same stimuli and a similar paradigm, Experiment 2.3 in return showed that in contrast to temporal ensemble statistics, spatial ensemble perception extracts the higher level gist of the faces: the ensemble representation of spatially presented unattractive faces is as unattractive as its components (symmetrical adaptation aftereffects between the attractive and unattractive face streams). Temporal and spatial ensemble representations, therefore, arise from distinct neural operations that produce qualitatively different perceptual outcomes.

This thesis, for the first time, revealed the computational mechanisms of ensemble statistics. The different mechanisms also reflect distinct neural encoding stages of the properties that make up facial attractiveness and reveal the hierarchical nature of the human face perception system. The ensemble coding of faces is interdependent with face processing. Therefore, it is reasonable to assume that all levels of ensemble statistics are nested with the associated visual processing.

In the temporal ensemble coding experiments, the faces were always presented at fast speed. Would such speed generate some morphing in itself? McKeeff et al. (2007) used fMRI and showed that even at a high speed (roughly 40 Hz, very similar to our high-speed RSVP), the V1 cortex activates more than its activation for the low speed face stream. This suggests that even at high speeds, the cortical visual processing regions are still capable of (at least partially) separating the visual inputs. And thus, the faces at high-speed RSVPs would not generate a morphing effect themselves, but they can be perceived as a stream by our visual system.

Although this thesis did not directly examine the combining effect of spatial and temporal ensemble coding, it is reasonable to believe (based on our findings in this thesis) that inner representation of the stimuli would be the 'gist' average of the 'computational' average of each temporal stream. For instance, if there are several streams of unattractive faces being presented in front of us, our visual system would tend to average the individual stream via 'computational' averaging. Thus, the visual system perceives several 'computational' averages all over the space. Therefore, our visual system would further average the 'computational' averages via 'gist' averaging: the several streams of unattractive faces would be perceived as mediocre to look at and would generate a similar adaptation aftereffect like 'UNA' conditions in Experiment 2.1 and 2.2. Future research should be conducted to directly study our hypothesis here.

What could impact the ensemble statistics of faces? As the result of Study 1 shows variance could not alter the adaptation aftereffects induced by a stream of faces. Therefore, the ensemble perception is (at least partly) immune to differences in variance. However, using direct report and adaptation paradigms, converging evidence from Study 4 showed both top-down and bottom-up attention heavily influence the ensemble statistics of facial expressions by changing the weight of each component. Thus, ensemble coding is a weighted average modulated by attention.

Would the ensemble coding of facial attractiveness also be immune to temporal frequency manipulation? It is reasonable to believe that the temporal frequency would also barely impact the 'computational averaging' of the facial attractiveness of the temporally presented faces, just like what has been observed in Experiment 1.3. Because of the overlapping neural mechanisms between the perception facial expression and that of facial attractiveness (Haxby & Gobbini, 2001; O'Doherty et al., 2003; Todorov et al., 2008), the temporal ensemble coding might be conducted by a similar computational mechanism and is largely immune to the temporal frequency control. Future research should directly examine the possible impact of temporal frequency on temporal ensemble coding of facial attractiveness as well as other facial characteristics.

In Study 4, the data suggests that the ensemble coding of facial expressions is similar to ensemble coding of size, as both of them are subject to attention manipulation. However, there is still a distinction among different levels of ensemble statistics. Haberman and colleagues (2015) directly compared the ensemble coding of faces against the ensemble coding of low-level stimuli. If all ensemble statistics are formed by the same mechanism, then the error in low-level ensemble should correlates with not only other low-level ensembles but also the high-level ones. However, they found the averaging error in the ensemble coding of facial expressions could be predicted by that of facial identity but not by those of size and orientation.

Thus, the data clearly shows a distinction between high- and low-level ensemble coding. As suggested by Whitney and Leib (2017), the ensemble coding of faces and the ensemble coding of low-level features (e.g. size, contrast) are inter-dependent, because the ensemble coding may reflect some general features of visual processing, rather than high- or low-level specific visual processing.

Finally, two lines of findings here hinted at the possible neural mechanisms for ensemble statistics of faces. Based on the findings of adaptation (Fox & Barton, 2007; Webster, 2011), the aftereffects results suggested that the ensemble statistics occur in the regions employed for face processing. This converges with the previous finding that the ensemble statistics of higher-level objects are distinctive from that of lower-level objects (Haberman et al., 2015). Moreover, the impact of attention on ensemble coding further suggested that the neural systems which code the ensemble of faces should be subject to attention. Taken together, these two lines of findings indicated the ensemble statistics of faces might occur at the core region for face processing, namely the OFA, STS, and FFA (Haxby & Gobbini, 2011; Haxby et al., 2000).

6.1.2. Mechanism of the Perception of Facial Expression

Previous research extensively studied the perception of facial expression, but we still do not know what happens when we see a group of faces, which is very common in daily life. Studies in RSVP suggest that the visual system is capable to correctly recalling certain items from the stream (Keysers & Perrett, 2002; Keysers, Xiao, Földiák, & Perrett, 2001; Potter, 1975, 2014). However, evidence from other studies indicated that the visual system cannot recall all of them but instead averages them together (Brady & Alvarez, 2015; Haberman et al., 2009; Haberman & Whitney, 2012; Nieuwenstein & Potter, 2006). Besides, most of the studies mentioned above employed the voluntary report (e.g. pick the test face that represents the mean emotion of the previous group). However, the adaptation paradigm we employed throughout this thesis is capable of examining involuntarily ensemble coding. The mechanism of the involuntary perception of facial expressions streams is less clear.

In the first study, an adaptation paradigm was used to probe the perception of temporally presented facial expressions. During the whole study, subjects were not instructed to respond or process the adapting face streams. Therefore, the induced aftereffects purely reflected the involuntary processing of the faces. The evidence from the first two experiments showed that the adaptation aftereffects of RSVP of facial expressions are comparable and correlated with those induced by the computer-generated average faces. Thus, the finding supported the notion that the visual system is capable of representing the temporally presented faces by ensemble coding.

Apparently, the stimuli used in Experiment 1.1 & 1.2 have small variance, while the faces in the real-life scenarios are distinctive in their facial expressions. Temporal frequencies and emotion variances were carefully induced in Experiment 1.3. As long as the mean emotion is the same, regardless of variances, those face streams evoked the similar adaptation aftereffects. Further controls in changing the facial identity (Experiment 1.4) closely resembled the findings. Therefore, the visual system processes the emotions of the stream of faces by involuntary ensemble statistics.

6.1.3. Mechanism of Facial Attractiveness Perception

How do past experiences with faces alter individuals' future judgments on attractiveness? In Experiment 3.1 & 3.2, results showed that a face was judged as most attractive after the viewer had recently experienced groups of unattractive faces, and *vice versa*. The adaptation aftereffects were correlated with the underlying mean attractiveness ratings for the groups of faces, suggesting that visual system implicitly averages the past experience of facial attractiveness and recalibrates the current perception.

When subjects judged a target face which was presented within a group of faces that varied in levels of attractiveness, the results in Experiment 3.3 & 3.4 showed that the ‘cheerleader/friend effect’ was indeed shaped by the attractiveness of the company the target face kept: the target appeared more attractive when surrounded by increasingly unattractive ‘friends’. The size of the friend effect was correlated with the mean attractiveness ratings for the surrounding group. However, even surrounded by the most attractive faces, unlike the adaptation aftereffect in Experiment 3.1 & 3.2, there was still a significant residual attractiveness boost. A further control in Experiment 3.4 suggested that the ‘social positive effect’ is not necessarily induced by being surrounded by other faces. When the target face was surrounded by six copies of itself, there was no significant attractiveness boost. Therefore, the results here suggested that the contrastive effect (similar to adaptation) between the ensemble perception of the different surrounding faces and the central target face determines the friend effect.

The results from these four experiments together suggested that facial attractiveness, in the real world, is dynamically perceived through implicit comparisons based upon its prior, and current, experiences. Also, the results delineated the possible mechanisms of the friend effect. Thus, this study bridged two key phenomena in facial attractiveness, namely adaptation and the cheerleader effect: they are both shaped and linked by ensemble perception. This study even provided the general public with a useful ‘tip’: if an individual wants to maximize his/her appearance, he/she should surround him/herself with unattractive friends.

Facial attractiveness could be affected by both low-level local information and high-level holistic information (Little et al., 2011). For example, skin tone is comparatively low-level local information, while the averageness is a kind of high-level holistic information determined by the second-order relationships among facial features (Maurer et al., 2002). In Experiment 2.1 and 2.2, the testing faces were at 75% size of the adaptors to minimize the local adaptation, and in Experiment 2.3 the adaptors and testing faces were presented at different locations. Therefore, the observed adaptation aftereffects in this thesis could not be fully explained by the local features. Moreover, using inverted faces, Olson and Marshuetz (2005) showed that the inverted attractive faces could not generate similar priming effects when they were upright. They further suggested that the perception of attractiveness requires the holistic processing. On the other hand, ensemble coding is widely judged as holistical processing (Whitney & Leib, 2017). For instance, Leib et al., (2012) found that the participants could report the mean of upright faces better than inverted ones. Taken these findings together, it is reasonable to assume that the ensemble coding of facial attractiveness is driven by the high-level information in a holistic manner. However, it is still worthwhile for future researchers to investigate the possible influence of low-level local information on the high-level facial attractiveness perception. For instance, the face composite paradigm (Rossion, 2013) and part-whole paradigm (Tanaka & Farah, 1993) should be used to validate the hypothesis.

Noticeably, this thesis examined the ensemble coding of facial expressions and facial attractiveness separately. The relationship between these two remains as an open question. Based on the current understanding of facial attractiveness, it is reasonable for one to believe there might be an interaction between ensemble coding of facial expressions and facial attractiveness. A growing body of evidence suggests that our perception of facial characteristics (e.g. attractiveness, trustworthiness) might be a consequence of overgeneralization of facial expressions (e.g. Todorov et al., 2008). O'Doherty and colleagues (2003) also showed that the orbital frontal cortex (OFC) could be activated by both attractive faces and face smile; suggesting a shared neural mechanism for perception of facial expression and facial attractiveness. For instance, a happy face should be perceived as more attractive to look at than a sad face, because the happiness in the face may further activate the neural systems evaluating facial attractiveness. Although we did not directly test this interaction hypothesis, however, we carefully picked neutral faces for Study 2 & 3, when testing the ensemble coding of facial attractiveness, to minimize the potential influence of facial expression. Future researchers should directly test the possible interaction of facial expression and attractiveness in ensemble coding to further clarify the mechanism behind ensemble coding.

6.1.4. Formation of Face Space

A substantial body of evidence from psychophysics and neural imaging studies suggested that the face processing network represents perceived faces using ‘face space’ (Bruce & Young, 2012; Leopold et al., 2001; Loffler et al., 2005; Nishimura et al., 2008; Oosterhof & Todorov, 2008; Rhodes et al., 2005; Sutherland et al., 2013; Todorov et al., 2015; Webster, 2011; Valentine, 1991, 2001; but see Jenkins et al., 2011). This theory, inspired by the face distinctiveness effect (Valentine, 2001), suggests that each face is uniquely represented in a multidimensional psychological space against the central norm. Thus, some researchers named this processing as norm-based coding of the face. Consequently, the centroid of the face space is the peak of the typicality: the center norm should be the most averaged face.

How are the face space and the face norm formed? Research into the other race effect looks into the possible mechanism. As Valentine (1991) suggested, the insufficient experience of faces from other races lead to a face space and norm which fail to represent the foreign faces adequately. For example, using an adaptation paradigm, Webster and colleagues (2004) found that Japanese students who are not familiar with Caucasians have different perceptual patterns in response to the Caucasian faces compared to native subjects. However, with an immersive experience abroad, those Japanese students' face perception of Caucasian faces got improved. Recently, some researchers reported that even four-year-old children are able to conduct certain ensemble processing (Sweeny, Wurnitsch, Gopnik, & Whitney, 2015), and could use norm-based coding for faces (Jeffery, Read, & Rhodes, 2013). Here, four studies together showed that the past experience of faces would update the face norm and bias the perception of the subsequently viewed faces. Studies in this thesis, together with those studies (see also, Rhodes, Neumann, Ewing, & Palermo, 2015), highlight a potential link between ensemble and norm-based face processing. Current findings delineated the formation of the facial expression norm during adaptation: our vision system implicitly integrates the multiple faces we encounter over time to the average face of the face stream. Consequently, the recent visual experience could update the norm of the face space shaped by ensemble statistics. Therefore, this thesis, for the first time, offers new insight into the updating procedure of the face space and face norm.

6.1.5. A New Paradigm to Study Ensemble Statistics

In most studies of ensemble statistics, the researchers required subjects to directly and voluntarily compute the mean of the stimuli. Some examples of this include choosing whether the testing face is more emotional than the mean of the preceding faces (e.g. Haberman et al., 2009; Haberman & Whitney, 2009), or choosing which side (of the screen) contains the circles with larger mean size (e.g. Chong & Treisman, 2003; Gorea, Belkoura, & Solomon, 2014); or choosing a face that matches the mean identity of the preceding faces (e.g. Haberman et al., 2015; Leib et al., 2014). The conventional measurements might only probe the ensemble representation in working memory, but not necessarily the encoding stage in the face processing regions.

Quite differently, this thesis tested the involuntary ensemble coding of faces by a visual adaptation paradigm (Webster et al., 2004; Ying & Xu, 2017). Being described as a ‘microelectrode for psychologists’, adaptation has been widely used in studying face perception (Frisby, 1979). Are adaptation aftereffects as capable to examine the ensemble statistics as conventional methods are? Study 2 compares the results between the adaptation paradigm and the direct report (similar to Haberman et al., 2009). The significant correlation between the reported mean attractiveness and the adaptation aftereffect suggests that the adaptation paradigm is capable of detecting the ensemble statistics. Also, the converging evidence from Study 4 using different measurements indicates the effectiveness and robustness of the adaptation paradigm in detecting the ensemble coding. Consequently, adaptation employed here is a powerful and reliable paradigm to study the mechanisms of involuntary ensemble statistics.

6.1.6. Implications for the Face Adaptation Aftereffect

Conventionally, experiments using a face adaptation paradigm always put the adaptors and the testing face at the same location to maximize the aftereffect (Leopold et al., 2001; Rhodes et al., 2003; Webster et al., 2004; Xu et al., 2008; Ying & Xu, 2017). After doing so, however, it is hard to discriminate the sources of the face adaptation aftereffect (Adams, Gray, Garner, & Graf, 2010; Afraz & Cavanagh, 2008; Leopold et al., 2001). Since face processing is hierarchical, adaptation to low-level features is possible to generate an aftereffect at higher-level face perception (Xu et al., 2008). To address this issue, some researchers shrunk the size of the test face (e.g. testing stimuli at 75% size of the adaptors) to counterbalance retinotopic adaptation (i.e. Rhodes et al., 2007; Burton et al., 2015; Burton et al., 2015; Zhao & Chubb, 2001), while some researchers floated the adaptor to avoid local adaptation (i.e., Bi, Su, Chen, & Fang, 2009; in this study, the adaptor moved slowly within a predefined area at the center of the screen). However, the residual overlapping between the adaptors and testing faces leaves an opening for low-level adaptation.

In Experiment 3.1 and 4.2, the adaptors and testing faces were presented at non-overlapping spatial locations, while adaptors appeared for brief durations (1s for facial attractiveness adaptation, and 200ms for facial expression adaptation). Such controls are typically thought to counteract low-level retinotopic effects. Thus, the observed aftereffects in these experiments could only be attributed to high-level face adaptation aftereffects. These experiments also offer new insights into face perception as well as high-level face adaptation.

6.1.7. The Face Databases

In this thesis, three face databases were used. Study 1 and Study 4 used the KDEF database to study the ensemble coding of facial expressions. Study 2 used the N-FEE face database to examine the different mechanisms of temporal and spatial ensemble statistics of facial attractiveness. Study 3 used the Oslo face database to study the relationship among different ensemble coding, the facial attractiveness adaptation aftereffect, and the friend effect. Both KDEF database and Oslo face database contain pictures from Caucasian models, while the N-FEE face database consists of pictures of local Singaporeans.

For the facial expression related studies, we used the KDEF database. The KDEF database is one of the most popular face databases among visual scientists and psychologists. It contains more than 4900 high quality images with excellent facial expressions and has been used in more than 2000 published papers (according to KDEF database). The quality of these facial expressions has been wildly acknowledged. We are aware that the modals in this dataset are all Caucasian; however, we do not believe this will affect the interpretation of the data. Although there is an ongoing debate as to whether the basic facial expressions are universal or not (Ekman, 1993; Jack, Garrod, & Schyns, 2014; Jack & Schyns, 2017), researchers still have consensus that the happy and sad expressions (the ones we used in our studies) are from distinctive emotion categories and can be perceived universally. Moreover, the emotional expressions from the KDEF database has been validated across different cultures (e.g. Goeleven et al., 2008; Yan, Andrews, & Young, 2016). Therefore, using the KDEF database allows us to examine the perception of emotional expression in a conventional way.

For the facial attractiveness related studies, we used the N-FEE face database (Study 2, Chapter III) and the Oslo face database (Study 3, Chapter IV). The N-FEE face database contains local Singaporean faces, while the Oslo face database consists of Caucasian faces. Previous researchers have shown that, at least in female attractiveness (what we studied in this thesis), is universal (regardless of viewer's ethnicity or gender). When studying the innate feature of female facial attractiveness, Perrett and colleagues (1994) found '…Caucasian and Japanese subjects showed the same pattern of preference'. The judgments of attractiveness towards female faces are almost perfectly correlated irrespective of the race being judged or judging (i.e., $r > .9$ in Cunningham, Roberts, Barbee, Druen, & Wu, 1995; Perrett et al., 1998; Rhodes et al., 2001). Moreover, the results from Experiment 2.3 (using local Singaporean faces) and Experiment 3.1 (using Caucasian faces) showed that both of the databases can generate similar adaptation aftereffects; meaning they are equally capable for our research. To minimize the possible unforeseeable influence of using different databases, all the stimuli used in Study 2 & 3 were selected using the judgements of our participants: for instance, the attractive faces were specifically picked based on participants' ratings, so that each attractive face is appreciated by everyone (taking part the study). Therefore, we have good reasons to believe that the databases we used in the studies should not affect the interpretation of our findings.

6.2. Limitation and Future Direction

The current thesis uses adaptation paradigms to study the mechanisms of spatial and temporal ensemble statistics of facial expression and attractiveness. There are, inevitably, some limitations in this thesis. This section discusses the limitations of this thesis, and also points to some uncharted territories in ensemble statistics research.

Firstly, the categories of the facial expressions were limited in the current work. In study 1, only happiness and sadness expressions were inspected. However, some other researchers in ensemble statistics used a more comprehensive selection of facial expressions. For instance, Elias and colleagues (2017) used morphed faces with expressions including angry, fearful, and happy to study how synchronization of expressions affect the ensemble perception. Noticeably, Experiment 4.2 used a happy and anger pair and replicated the findings in other experiments with happy and sad faces. Still, testing ensemble statistics with other facial expressions would make the current findings even more convincing. Moreover, often neglected by researchers, using non-emotional expressions may clarify whether ensemble coding is an innate ability or merely reflects the expertise of the subjects and their experiences of facial expressions. Therefore, future experiments may apply an increased repertoire of categories of facial expressions to study the mechanism behind ensemble statistics further.

Secondly, the range of facial attractiveness was limited in Study 2 & 3. The face databases used in these two studies both contain pictures taken in well-controlled laboratory settings. However, since they both consist of ordinary people without makeup, there are no extremely attractive faces as appealing as models or actresses in both datasets. As discussed in Study 3, using extremely attractive faces may entirely abolish the friend effect (by fully canceling out the social positive effect), and further clarifying the mechanism of the friend effect. Therefore, future research may use models or actresses to further examine the ensemble statistics of facial attractiveness. Also, future research should also consider using faces with makeup to mimic the real-life scenarios.

Thirdly, the spatial arrangements of the adaptors and the test faces in Experiment 2.3 as well as that in Experiment 3.1 do not fully represent all kinds of spatial ensemble coding. Despite this, we believe the distinguished representations of spatial and temporal ensemble coding (Study 2) were not a consequence of the differences between the adaptor-test-overlap (Experiment 2.1 & 2.2) and adaptor-test-nonoverlap (Experiment 2.3). However, testing Experiment 2.3 with other designs could further clarify the computational mechanisms of spatial ensemble coding. Therefore, it is worthwhile for future researchers to evaluate the impact of the spatial arrangements of adaptors and test faces on adaptation as well as ensemble coding.

Fourthly, whether the temporal frequency of RSVP streams can affect the temporal ensemble coding of facial attractiveness is unclear. Although the results in Experiment 1.3 suggest that the temporal frequency can hardly affect the ensemble representation of facial expressions, we cannot assert that the same pattern would definitely be observed in facial attractiveness, considering the different neural mechanisms behind them. Future researchers should further clarify the impact of different temporal frequencies on the temporal ensemble coding of facial attractiveness.

Lastly, although the adaptation paradigm is a powerful tool to study ensemble statistics, it is not omnipotent at scrutinizing every facet of ensemble statistics. A growing body of evidence recently showed that the human vision system uses both central tendency and dispersion (as the name of ‘ensemble statistics' implies) to form the ensemble representation of the visual inputs (Whitney & Leib, 2017). Noticeably, only the central tendency part was examined in this thesis. However, the current adaptation paradigm in this thesis is not capable to fully examining the mechanisms behind the extraction of dispersion. On the other hand, the adaptation paradigm requires lots of repetition to estimate the aftereffect; it is naturally incapable of reflecting the changes between trials. Thus, the current method is unable to examine the ‘subsampling' hypothesis which could only be found in a trial-by-trial analysis. Multiple repetitions may perceptually train subjects to be better at ensemble statistics, so that the observed data might have been exaggerating the ensemble perception. Consequently, future research in ensemble statistics should utilize other measurements. Moreover, considering the limited number of adaptors (in spatial ensemble coding experiments), it is possible that the participants memorized the adaptors. In future research in adaptation, more adaptors should be used to counterbalance such limitations.

The future researchers shall examine ensemble statistics in other facial characteristics, such as trustworthiness, which is also fundamental to social communication (Oosterhof & Todorov, 2008; Sutherland et al., 2013). Besides, could the facial symmetry level and the sexual dimorphism, which are essential parts of facial attractiveness as averageness, be perceived involuntarily by ensemble coding in our visual system is also an important question for future researchers.

6.3. Conclusion

In conclusion, studies from this thesis systematically examined the mechanisms behind the spatial and temporal ensemble statistics of facial expressions and attractiveness. They linked the ensemble statistics with essential phenomena in face perception.

There were several lines of findings here. Firstly, the studies all suggested that the human visual system is able to involuntarily form the ensemble representation of facial information across time and space. Secondly, the ensemble statistics of temporally presented faces are more likely to involve a low-level ‘computational' averaging, while that of spatially presented faces favor a high-level gist averaging. Thirdly, face perception is heavily influenced by the present and the previous experience of faces by the ensemble statistics. Lastly, the ensemble perception of faces is not a simple arithmetic mean, but a weighted average (altered by attention).

These findings have psychophysical, neural and social implications. The results shed light on a refined theoretical framework of face perception in the group context.

References

Abbas, Z.A., Duchaine, B., & Duchaineô, B. (2008). The Role of Holistic Processing in Judgments of Facial Attractiveness. *Perception, 37*, 1187-1196. doi:10.1068/p5984

Adams, W. J., Gray, K. L. H., Garner, M., & Graf, E. W. (2010). High-Level Face Adaptation Without Awareness. *Psychological Science, 21*, 205-210. doi:10.1177/0956797609359508

Adolphs, R. (2006). Perception and emotion: How we recognize facial expressions. *Current Directions in Psychological Science, 15*(5), 222-226. doi:DOI 10.1111/j.1467-8721.2006.00440.x

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433*(7021), 68-72. doi:10.1038/nature03086

Afraz, S. R., & Cavanagh, P. (2008). Retinotopy of the face aftereffect. *Vision Research, 48*, 42-54. doi:10.1016/j.visres.2007.10.028

Allman, J., Miezin, F., & McGuinness, E. (1985). Direction-and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception, 14*(2), 105-126.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences, 15*, 122-131. doi:10.1016/j.tics.2011.01.003

Alvarez, G. A., & Oliva, A. (2008). The Representation of Simple Ensemble Visual Features Outside the Focus of Attention. *Psychological Science, 19*, 392-398. doi:10.1111/j.1467-9280.2008.02098.x

Alvarez, G. A., & Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proceedings of the National Academy of Sciences, 106*, 7345-7350. doi:10.1073/pnas.0808981106

Anderson, N. H., Lindner, R., & Lopes, L. L. (1973). Integration theory applied to judgments of group attractiveness. *Journal of Personality and Social Psychology, 26*, 400-408. doi:10.1037/h0034441

Andrews, T. J., & Ewbank, M. P. (2004). Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *Neuroimage, 23*, 905-913. doi:10.1016/j.neuroimage.2004.07.060

Anzures, G., Mondloch, C. J., & Lackner, C. (2009). Face Adaptation and Attractiveness Aftereffects in 8-Year-Olds and Adults. *Child Development, 80*, 178-191. doi:10.1111/j.1467-8624.2008.01253.x

Ariely, D. (2001). Seeing Sets: Representation by Statistical Properties. *Psychological Science, 12*, 157-162. doi:10.1111/1467-9280.00327

Awh, E., Vogel, E. K., & Oh, S.-H. H. (2006). Interactions between attention and working memory. *Neuroscience, 139*, 201-208. doi:10.1016/j.neuroscience.2005.08.023

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated Measures Correlation. *Frontiers in Psychology, 8*. doi:10.3389/fpsyg.2017.00456

Barraclough, N. E., & Perrett, D. I. (2011). From single cells to social perception. Philosophical *Transactions of the Royal Society of London B: Biological Sciences, 366*(1571), 1739-1752.

Bartolomeo, P., Vuilleumier, P., & Behrmann, M. (2015). The whole is greater than the sum of the parts: Distributed circuits in visual cognition. *Cortex, 72*, 1-4. doi:10.1016/j.cortex.2015.09.001

Behrmann, M., & Plaut, D. C. (2013). Distributed circuits, not circumscribed centers, mediate visual recognition (vol 17, pg 210, 2013). *Trends in Cognitive Sciences, 17*, 361. doi:10.1016/j.tics.2013.05.009

Bi, T. Y., Su, J. Z., Chen, J., & Fang, F. (2009). The role of gaze direction in face viewpoint aftereffect. *Vision Research, 49*(18), 2322-2327. doi:10.1016/j.visres.2009.07.002

Bisley, J. W., & Goldberg, M. E. (2003). Neuronal activity in the lateral intraparietal area and spatial attention. *Science, 299*(5603), 81-86. doi:DOI 10.1126/science.1077395

Brady, T. F., & Alvarez, G. A. (2011). Hierarchical Encoding in Visual Working Memory : Ensemble Statistics Bias Memory for Individual Items. doi:10.1177/0956797610397956

Brady, T. F., & Alvarez, G. A. (2015). No evidence for a fixed object limit in working memory: Spatial ensemble representations inflate estimates of working memory capacity for complex objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 921-929. doi:10.1037/xlm0000075

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433-436. doi:10.1163/156856897x00357

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77*, 305-327. doi:10.1111/j.2044-8295.1986.tb02199.x

Buckingham, G., DeBruine, L. M., Little, A. C., Welling, L. L. M., Conway, C. A., Tiddeman, B. P., & Jones, B. C. (2006). Visual adaptation to masculine and feminine faces influences generalized preferences and perceptions of trustworthiness. *Evolution and Human Behavior, 27*, 381-389. doi:10.1016/j.evolhumbehav.2006.03.001

Burr, D., & Ross, J. (2008). A visual sense of number. *Current Biology, 18*(6), 425-428. doi:10.1016/j.cub.2008.02.052

Burt, D. M., & Perrett, D. I. (1997). Perceptual asymmetries in judgements of facial attractiveness, age, gender, speech and expression. *Neuropsychologia, 35*(5), 685-693. doi:Doi 10.1016/S0028-3932(96)00111-X

Burton, N., Jeffery, L., Calder, A., & Rhodes, G. (2014). Adaptation to an average expression improves discrimination of facial expressions. *Journal of Vision, 14*, 815. doi:10.1167/14.10.815

Burton, N., Jeffery, L., Calder, A. J., & Rhodes, G. (2015). How is facial expression coded? *Journal of Vision, 15*, 1. doi:10.1167/15.1.1

Calder, A. J., Beaver, J. D., Davis, M. H., Van Ditzhuijzen, J., Keane, J., & Lawrence, A. D. (2007). Disgust sensitivity predicts the insula and pallidal response to pictures of disgusting foods. *European Journal of Neuroscience, 25*, 3422-3428. doi:10.1111/j.1460-9568.2007.05604.x

Calder, A. J., Jenkins, R., Cassel, A., & Clifford, C. W. G. (2008). Visual representation of eye gaze is coded by a nonopponent multichannel system. *Journal of Experimental Psychology: General, 137*, 244-261. doi:10.1037/0096-3445.137.2.244

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience, 6*, 641-651. doi:10.1038/nrn1724

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 527-551. doi:10.1037/0096-1523.26.2.527

Campbell, J., & Burke, D. (2009). Evidence that identity-dependent and identity-independent neural populations are recruited in the perception of five basic emotional facial expressions. *Vision Research, 49*, 1532-1540. doi:10.1016/j.visres.2009.03.009

Chelnokova, O., Laeng, B., Eikemo, M., Riegels, J., Loseth, G., Maurud, H., . . . Leknes, S. (2014). Rewards of beauty: the opioid system mediates social motivation in humans (vol 19, pg 746, 2014). *Molecular Psychiatry, 19*(12), 1342-1342. doi:10.1038/mp.2014.149

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research, 43*, 393-404. doi:10.1016/s0042-6989(02)00596-5

Chong, S. C., & Treisman, A. (2005). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics, 67*(1), 1-13. doi:Doi 10.3758/Bf03195009

Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the Bandwidth of Perceptual Experience? *Trends in Cognitive Sciences, 20*(5), 324-335. doi:10.1016/j.tics.2016.03.006

Collins, J. A., & Olson, I. R. (2014). Beyond the FFA: The role of the ventral anterior temporal lobes in face processing. *Neuropsychologia, 61*, 65-79. doi:10.1016/j.neuropsychologia.2014.06.005

Corbett, J. E., Wurnitsch, N., Whitney, D., Schwartz, A., & Whitney, D. (2012). An aftereffect of adaptation to mean size. *Visual Cognition, 20*, 211-231. doi:10.1080/13506285.2012.657261

Cowan, N. (2010). The Magical Mystery Four. *Current Directions in Psychological Science, 19*, 51-57. doi:10.1177/0963721409359277

Cunningham, M. R., Roberts, A. R., Barbee, A. P., Druen, P. B., & al, e. (1995). "Their ideas of beauty are, on the whole, the same as ours": Consistency and variability in the cross-cultural perception of female physical attractiveness. *Journal of Personality and Social Psychology, 68*, 261-279. doi:10.1037//0022-3514.68.2.261

de Fockert, J. W., & Marchant, A. P. (2008). Attention modulates set representation by statistical properties. *Perception & Psychophysics, 70*, 789-794. doi:10.3758/pp.70.5.789

DeBruine, L., & Tiddeman, B. (2017). *WebMorph.*, Retrieved from http://webmorph.org.

DeBruine, L. M., Jones, B. C., Unger, L., Little, A. C., & Feinberg, D. R. (2007). Dissociating averageness and attractiveness: Attractive faces are not always average. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1420-1430. doi:10.1037/0096-1523.33.6.1420

Deffenbacher, K. A., Vetter, T., Johanson, J., & O'Toole, A. J. (1998). Facial Aging, Attractiveness, and Distinctiveness. *Perception, 27*, 1233-1243. doi:10.1068/p271233

Desimone, R., Albright, T. D., Gross, C. G., Bruce, C. (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci 4*:2051–2062.

Desimone, R. (1991) Face-selective cells in the temporal cortex of monkeys. J *Cogn Neurosci 3*:1– 8.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5*. doi:ARTN 78110.3389/fpsyg.2014.00781

Duchaine, B., & Yovel, G. (2015). A Revised Neural Framework for Face Processing. *Annual Review of Vision Science, Vol 1, 1*, 393-416. doi:10.1146/annurev-vision-082114-035518

Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport, 11*(10), 2319-2324.

Ekman, P. (1993). Facial expression and emotion. *American Psychologist, 48*, 384-392. doi:10.1037//0003-066x.48.4.384

Elias, E., Dyer, M., & Sweeny, T. D. (2017). Ensemble Perception of Dynamic Emotional Groups. *Psychological Science, 28*, 193-203. doi:10.1177/0956797616678188

Engell, A. D., Todorov, A., & Haxby, J. V. (2010). Common neural mechanisms for the evaluation of facial trustworthiness and emotional expressions as revealed by behavioral adaptation. *Perception, 39*(7), 931-941. doi:10.1068/p6633

Ewing, L., Rhodes, G., & Pellicano, E. (2010). Have you got the look? Gaze direction affects judgements of facial attractiveness. *Visual Cognition, 18*(3), 321-330. doi:10.1080/13506280902965599

Fang, F., Boyaci, H., Kersten, D., & Murray, S. O. (2008). Attention-Dependent Representation of a Size Illusion in Human V1. *Current Biology, 18*(21), 1707-1712. doi:10.1016/j.cub.2008.09.025

Fang, F., & He, S. (2005). Viewer-centered object representation in the human visual system revealed by viewpoint aftereffects. *Neuron, 45*(5), 793-800. doi:10.1016/j.neuron.2005.01.037

Finkbeiner, M., & Palermo, R. (2009). The Role of Spatial Attention in Nonconscious Processing. *Psychological Science, 20*, 42-51. doi:10.1111/j.1467-9280.2008.02256.x

Fischer, J., & Whitney, D. (2011). Object-level visual information gets through the bottleneck of crowding. *Journal of Neurophysiology, 106*, 1389-1398. doi:10.1152/jn.00904.2010

Fisher, C. I., Hahn, A. C., DeBruine, L. M., & Jones, B. C. (2014). Integrating shape cues of adiposity and color information when judging facial health and attractiveness. *43*, 499-508. doi:10.1068/p7728

Fox, C. J., & Barton, J. J. S. (2007). What is adapted in face adaptation? The neural representations of expression in the human visual system. *Brain Research, 1127*, 80-89. doi:10.1016/j.brainres.2006.09.104

Fox, C. J., Oruc, I., & Barton, J. J. S. (2008). It doesn't matter how you feel. The facial identity aftereffect is invariant to changes in facial expression. *Journal of Vision, 8*(3). doi:Artn 1110.1167/8.3.11

Furl, N. (2016). Facial-Attractiveness Choices Are Predicted by Divisive Normalization. *Psychological Science, 27*, 1379-1387. doi:10.1177/0956797616661523

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level (vol 12, pg 499, 2000). *Journal of Cognitive Neuroscience, 12*(5), 912-912.

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: bridging selective attention and working memory. *Trends in Cognitive Sciences, 16*, 129-135. doi:10.1016/j.tics.2011.11.014

Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *International journal of endocrinology and metabolism, 10*(2), 486.

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia, 45*, 32-41. doi:10.1016/j.neuropsychologia.2006.04.015

Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska directed emotional faces: a validation study. *Cognition and emotion*, *22*(6), 1094-1118.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Research, 41*(17), 2261-2271. doi:Doi 10.1016/S0042-6989(01)00097-9

Grammer, K., & Thornhill, R. (1994). Human (Homo sapiens) facial attractiveness and sexual selection: The role of symmetry and averageness. *Journal of Comparative Psychology, 108*, 233-242. doi:10.1037//0735-7036.108.3.233

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nat Neurosci, 7*(5), 555-562. doi:10.1038/nn1224

Gross, C. G., Rocha-Miranda, C. E., Bender, D. B. (1972) Visual properties of neurons in inferotemporal cortex of the macaque. *J Neurophysiol 35*:96 –111.

Haberman, J., Brady, T. F., & Alvarez, G. A. (2015). Individual differences in ensemble perception reveal multiple, independent levels of ensemble representation. *Journal of Experimental Psychology: General, 144*, 432-446. doi:10.1037/xge0000053

Haberman, J., Harp, T., & Whitney, D. (2009). Averaging facial expression over time. *Journal of Vision, 9*, 1. doi:10.1167/9.11.1

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Curr Biol, 17*(17), R751-753. doi:10.1016/j.cub.2007.06.039

Haberman, J., & Whitney, D. (2009). Seeing the mean: Ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 718-734. doi:10.1037/a0013899

Haberman, J., & Whitney, D. (2010). The visual system discounts emotional deviants when extracting average expression. *Attention, Perception & Psychophysics, 72*, 1825-1838. doi:10.3758/app.72.7.1825

Haberman, J., & Whitney, D. (2012). Ensemble Perception:summarizing the scene and broadening the limits of visual processing. *From perception to consciousness: Searching with Anne Treisman, 339-349.*

Harris, R. J., Young, A. W., & Andrews, T. J. (2012). Morphing between expressions dissociates continuous from categorical representations of facial expression in the human brain. doi:10.1073/pnas.1212207110/

Haxby, J. V., & Gobbini, M. I. Distributed Neural Systems for Face Perception, Oxford University Press (2011).

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*(6), 223-233. doi:Doi 10.1016/S1364-6613(00)01482-0

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry, 51*(1), 59-67. doi:Doi 10.1016/S0006-3223(01)01330-0

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception.

Holmes, A., Vuilleumier, P., & Eimer, M. (2003). The processing of emotional facial expression is gated by spatial attention : evidence from event-related brain potentials. *16*, 174-184. doi:10.1016/S0926-6410(02)00268-9

Hönekopp, J. (2006). Once more: Is beauty in the eye of the beholder? Relative contributions of private and shared taste to judgments of facial attractiveness. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 199-209. doi:10.1037/0096-1523.32.2.199

Hsu, S. M. S.-m., & Young, A. W. (2007). Adaptation effects in facial expression recognition. *Visual Cognition, 11*, 871-899. doi:10.1080/13506280444000030

Humphreys, K., Avidan, G., & Behrmann, M. (2006). A detailed investigation of facial expression processing in congenital prosopagnosia as compared to acquired prosopagnosia. *Experimental Brain Research, 176*, 356-373. doi:10.1007/s00221-006-0621-5

Jack, R. E., Garrod, O. G., & Schyns, P. G. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Curr Biol, 24*, 187-192. doi:10.1016/j.cub.2013.11.064

Jack, R. E., & Schyns, P. G. (2017). Toward a Social Psychophysics of Face Communication. *Annual Review of Psychology, Vol 68, 68*, 269-297. doi:10.1146/annurev-psych-010416-044242

Jeffery, L., Read, A., & Rhodes, G. (2013). Four year-olds use norm-based coding for face identity. *Cognition, 127*, 258-263. doi:10.1016/j.cognition.2013.01.008

Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience, 6*(10), 766.

Jones, B. C., DeBruine, L. M., Little, A. C., Burriss, R. P., & Feinberg, D. R. (2007). Social transmission of face preferences among humans. *Proceedings of the Royal Society B: Biological Sciences, 274*, 899-903. doi:10.1098/rspb.2006.0205

Jones, B. C., DeBruine, L. M., Little, A. C., & Feinberg, D. R. (2007). The valence of experiences with faces influences generalized preferences. *Journal of Evolutionary Psychology, 5*, 119-129. doi:10.1556/jep.2007.1001

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

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B-Biological Sciences, 361*(1476), 2109-2128. doi:10.1098/rstb.2006.1934

Keysers, C., & Perrett, D. I. (2002). Visual masking and RSVP reveal neural competition. *Trends in Cognitive Sciences, 6*, 120-125. doi:10.1016/s1364-6613(00)01852-0

Keysers, C., Xiao, D.-K. K., Földiák, P., & Perrett, D. I. (2001). The Speed of Sight. *Journal of Cognitive Neuroscience, 13*, 90-101. doi:10.1162/089892901564199

Kleinhans, N. M., Richards, T., Johnson, L. C., Weaver, K. E., Greenson, J., Dawson, G., & Aylward, E. (2011). fMRI evidence of neural abnormalities in the subcortical face processing system in ASD. *Neuroimage, 54*(1), 697-704.

Langlois, J. H., & Roggman, L. A. (1990). Attractive Faces Are Only Average. *Psychological Science, 1*, 115-121. doi:DOI 10.1111/j.1467-9280.1990.tb00079.x

Lawson, R. P., Clifford, C. W. G., & Calder, A. J. (2011). A real head turner: Horizontal and vertical head directions are multichannel coded. *Journal of Vision, 11*, 17. doi:10.1167/11.9.17

Leib, A. Y., Puri, A. M., Fischer, J., Bentin, S., Whitney, D., & Robertson, L. (2012). Crowd perception in prosopagnosia. *Neuropsychologia, 50*, 1698-1707. doi:10.1016/j.neuropsychologia.2012.03.026

Leib, A., Fischer, J., Liu, Y., Qiu, S., Robertson, L., & Whitney, D. (2014). Ensemble crowd perception: A viewpoint-invariant mechanism to represent average crowd identity. *Journal of Vision, 14*, 26. doi:10.1167/14.8.26

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience, 4*, 89-94. doi:10.1038/82947

Leopold, D. A., & Rhodes, G. (2010). A comparative view of face perception. *Journal of Comparative Psychology, 124*, 233-251. doi:10.1037/a0019460

Leopold, D. A., Rhodes, G., Muller, K. M., & Jeffery, L. (2005). The dynamics of visual adaptation to faces. *Proceedings of the Royal Society B: Biological Sciences, 272*, 897-904. doi:10.1098/rspb.2004.3022

Li, K. A., & Yeh, S. L. (2017). Mean size estimation yields left-side bias: Role of attention on perceptual averaging. *Attention Perception & Psychophysics, 79*(8), 2538-2551. doi:10.3758/s13414-017-1409-3

Light, L. L., Kayrastuart, F., & Hollander, S. (1979). Recognition Memory for Typical and Unusual Faces. *Journal of Experimental Psychology-Human Learning and Memory, 5*(3), 212-228. doi:Doi 10.1037//0278-7393.5.3.212

Little, A. C., Burt, D. M., Penton-Voak, I. S., & Perrett, D. I. (2001). Self-perceived attractiveness influences human female preferences for sexual dimorphism and symmetry in male faces. *Proceedings of the Royal Society B: Biological Sciences, 268*, 39-44. doi:10.1098/rspb.2000.1327

Little, A. C., DeBruine, L. M., & Jones, B. C. (2005). Sex-contingent face after-effects suggest distinct neural populations code male and female faces. doi:10.1098/rspb.2005.3220

Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). Facial attractiveness: evolutionary based research. *Philos Trans R Soc Lond B Biol Sci, 366*, 1638-1659. doi:10.1098/rstb.2010.0404

Little, A. C., Jones, B. C., Waitt, C., Tiddeman, B. P., Feinberg, D. R., David, I., . . . Marlowe, F. W. (2008). Symmetry Is Related to Sexual Dimorphism in Faces : Data Across Culture and Species. *3*, 1-8. doi:10.1371/journal.pone.0002106

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: an MEG study. *Nat Neurosci, 5*(9), 910-916. doi:10.1038/nn909

Liu, J., Harris, A., & Kanwisher, N. (2010). Perception of face parts and face configurations: an FMRI study. *J Cogn Neurosci, 22*, 203-211. doi:10.1162/jocn.2009.21203

Lovejoy, L. P., & Krauzlis, R. J. (2010). Inactivation of primate superior colliculus impairs covert selection of signals for perceptual judgments. *Nature neuroscience, 13*(2), 261.

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of neurophysiology, 77*(1), 24-42.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska directed emotional faces (KDEF). *CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet*, 91-630.

Luo, A. X., & Zhou, G. (2018). Ensemble perception of facial attractiveness. *Journal of vision, 18*(8), 7-7.

Luo, C., Wang, Q., Schyns, P. G., Kingdom, F. A. A., & Xu, H. (2015). Facial Expression Aftereffect Revealed by Adaption to Emotion-Invisible Dynamic Bubbled Faces. *PLOS ONE, 10*, e0145877. doi:10.1371/journal.pone.0145877

Kay, K. N., Weiner, K. S., Grill-Spector, G. (2015), Attention reduces spatial uncertainty in human ventral temporal cortex. *Current Biology*, 25 (5), 595-600.

Maloney, L. T., & Dal Martello, M. F. (2006). Kin recognition and the perceived facial similarity of children. *Journal of Vision, 6*, 4. doi:10.1167/6.10.4

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in cognitive sciences, 6*(6), 255-260.

McKeeff, T. J., Remus, D. A., & Tong, F. (2007). Temporal limitations in object processing across the human ventral visual pathway. *Journal of Neurophysiology, 98*(1), 382-393. doi:10.1152/jn.00568.2006

Moran, J., & Desimone, R. (1985). Selective Attention Gates Visual Processing in the Extrastriate Cortex. *Science, 229*(4715), 782-784. doi:DOI 10.1126/science.4023713

Nguyen, M. N., Nishimaru, H., Matsumoto, J., Van Le, Q., Hori, E., Maior, R. S., ... & Nishijo, H. (2016). Population coding of facial information in the monkey superior colliculus and pulvinar. *Frontiers in neuroscience, 10*, 583.

Nieuwenstein, M. R., & Potter, M. C. (2006). Temporal limits of selection and memory encoding: A comparison of whole versus partial report in rapid serial visual presentation. *Psychol Sci, 17*, 471-475. doi:10.1111/j.1467-9280.2006.01730.x

O'Toole, A. J., Price, T., Vetter, T., Bartlett, J. C., & Blanz, V. (1999). 3D shape and 2D surface textures of human faces: the role of “averages” in attractiveness and age. *Image and Vision Computing, 18*, 9-19. doi:10.1016/s0262-8856(99)00012-8

O’Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M., & Dolan, R. J. (2003). Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia, 41*(2), 147-155. doi:10.1016/s0028-3932(02)00145-8

Olson, I. R., & Marshuetz, C. (2005). Facial attractiveness is appraised in a glance. *Emotion, 5*(4), 498.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, 105*, 11087-11092. doi:10.1073/pnas.0805664105

Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia, 45*(1), 75-92. doi:10.1016/j.neuropsychologia.2006.04.025

Pegors, T. K., Mattar, M. G., Bryan, P. B., & Epstein, R. A. (2015). Simultaneous perceptual and response biases on sequential face attractiveness judgments. *Journal of Experimental Psychology: General, 144*, 664-673. doi:10.1037/xge0000069

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*, 437-442. doi:10.1163/156856897x00366

Penton-Voak, I. S., Jones, B. C., Little, A. C., Baker, S., Tiddeman, B., Burt, D. M., & Perrett, D. I. (2001). Symmetry, sexual dimorphism in facial proportions and male facial attractiveness. *Proceedings of the Royal Society of London B: Biological Sciences, 268*(1476), 1617-1623.

Perrett, D. I., Burt, D. M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., & Edwards, R. (1999). Symmetry and Human Facial Attractiveness. *Evolution and Human Behavior, 20*, 295-307. doi:10.1016/s1090-5138(99)00014-8

Perrett, D. I., Hietanen, J. K., Oram, M. W., Benson, P. J., & Rolls, E. T. (1992). Organization and Functions of Cells Responsive to Faces in the Temporal Cortex [and Discussion]. *Philosophical Transactions of the Royal Society B: Biological Sciences, 335*, 23-30. doi:10.1098/rstb.1992.0003

Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., . . . Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature, 394*, 884-887. doi:10.1038/29772

Perrett, D. I., May, K. A., & Yoshikawa, S. (1994). Facial shape and judgements of female attractiveness. *Nature, 368*, 239-242. doi:10.1038/368239a0

Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. *Nature Reviews Neuroscience, 11*, 773-782. doi:10.1038/nrn2920

Pitcher, D., Walsh, V., Yovel, G., & Duchaine, B. (2007). TMS evidence for the involvement of the right occipital face area in early face processing. *Current Biology, 17*(18), 1568-1573. doi:10.1016/j.cub.2007.07.063

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3-25. doi:10.1080/00335558008248231

Potter, M. (1975). Meaning in visual search. *Science, 187*, 965-966. doi:10.1126/science.1145183

Potter, M. (2014). Detecting Meaning in Rapid Pictures. *Attention, Perception, & Psychophysics,*.

Raymond, J. E., Fenske, M. J., & Tavassoli, N. T. (2003). Selective attention determines emotional responses to novel visual stimuli. *Psychological Science, 14*(6), 537-542. doi:DOI 10.1046/j.0956-7976.2003.psci_1462.x

Rensink, R. A. (2000). Seeing, sensing, and scrutinizing. *Vision Research, 40*(10-12), 1469-1487. doi:Doi 10.1016/S0042-6989(00)00003-1

Rhodes, G. (2006). The Evolutionary Psychology of Facial Beauty. *Annual Review of Psychology, 57*, 199-226. doi:10.1146/annurev.psych.57.102904.190208

Rhodes, G., & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Research, 46*, 2977-2987. doi:10.1016/j.visres.2006.03.002

Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W., & Nakayama, K. (2003). Fitting the mind to the world: face adaptation and attractiveness aftereffects. *Psychol Sci, 14*, 558-566.

Rhodes, G., Maloney, L. T., Turner, J., & Ewing, L. (2007). Adaptive face coding and discrimination around the average face. *47*, 974-989. doi:10.1016/j.visres.2006.12.010

Rhodes, G., & Tremewan, T. (1996). Averageness, Exaggeration, and Facial Attractiveness. *Psychological Science, 7*, 105-110. doi:10.1111/j.1467-9280.1996.tb00338.x

Rhodes, G., Yoshikawa, S., Clark, A., Lee, K., McKay, R., & Akamatsu, S. (2001). Attractiveness of Facial Averageness and Symmetry in Non-Western Cultures: In Search of Biologically Based Standards of Beauty. *Perception, 30*, 611-625. doi:10.1068/p3123

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition, 21*(2), 139-253.

Rossion, B. (2014). Understanding individual face discrimination by means of fast periodic visual stimulation. *Experimental Brain Research, 232*, 1599-1621. doi:10.1007/s00221-014-3934-9

Rouder, J. N., Speckman, P. L., Sun, D. C., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*(2), 225-237. doi:10.3758/Pbr.16.2.225

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychol Bull, 115*(1), 102-141.

Schweinberger, S. R., Zäske, R., Walther, C., Golle, J., Kovács, G., & Wiese, H. (2010). Young without plastic surgery: Perceptual adaptation to the age of female and male faces. *Vision Research, 50*, 2570-2576. doi:10.1016/j.visres.2010.08.017

Schyns, P. G., Petro, L. S., & Smith, M. L. (2007). Dynamics of Visual Information Integration in the Brain for Categorizing Facial Expressions. *Current Biology, 17*, 1580-1585. doi:10.1016/j.cub.2007.08.048

Scott, S. K., Young, A. W., Calder, A. J., Hellawell, D. J., Aggleton, J. P., & Johnsons, M. (1997). Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature, 385*(6613), 254.

Seyama, J., & Nagayama, R. S. (2006). Eye direction aftereffect. *Psychological Research-Psychologische Forschung, 70*(1), 59-67. doi:10.1007/s00426-004-0188-3

Störmer, V. S., Alvarez, G. A., Stormer, V. S., & Alvarez, G. A. (2016). Attention Alters Perceived Attractiveness. *Psychol Sci, 27*, 563-571. doi:10.1177/0956797616630964

Susilo, T., McKone, E., & Edwards, M. (2010). What shape are the neural response functions underlying opponent coding in face space? A psychophysical investigation. *Vision Research, 50*, 300-314. doi:10.1016/j.visres.2009.11.016

Susilo, T., McKone, E., & Edwards, M. (2010). What shape are the neural response functions underlying opponent coding in face space? A psychophysical investigation. *Vision Research, 50*(3), 300-314. doi:10.1016/j.visres.2009.11.016

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition, 127*, 105-118. doi:10.1016/j.cognition.2012.12.001

Sweeny, T. D., & Whitney, D. (2014). Perceiving crowd attention: ensemble perception of a crowd's gaze. *Psychol Sci, 25*, 1903-1913. doi:10.1177/0956797614544510

Sweeny, T. D., Wurnitsch, N., Gopnik, A., & Whitney, D. (2014). Ensemble perception of size in 4-5-year-old children. *Developmental Science, 18*, 556-568. doi:10.1111/desc.12239

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology Section A, 46*(2), 225-245.

Tanaka, J. W., Kaiser, M. D., Butler, S., & Le Grand, R. (2012). Mixed emotions: Holistic and analytic perception of facial expressions. *Cognition & emotion, 26*(6), 961-977. doi: 10.1080/02699931.2011.630933

Thornhill, R., & Gangestad, S. W. (1999). Facial attractiveness. *Trends in Cognitive Sciences, 3*, 452-460. doi:10.1016/s1364-6613(99)01403-5

Tiddeman, B., Burt, M., & Perrett, D. (2001). Prototyping and transforming facial textures for perception research. *Ieee Computer Graphics and Applications, 21*(5), 42-50. doi:Doi 10.1109/38.946630

Todorov, A., & Engell, A. D. (2008). The role of the amygdala in implicit evaluation of emotionally neutral faces. *Social Cognitive and Affective Neuroscience, 3*, 303-312. doi:10.1093/scan/nsn033

Todorov, A., Gobbini, M. I., Evans, K. K., & Haxby, J. V. (2007). Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia, 45*, 163-173. doi:10.1016/j.neuropsychologia.2006.04.018

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in cognitive sciences, 12*(12), 455-460.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cogn Psychol, 12*(1), 97-136.

Tsotsos, J. K. (1990). Analyzing vision at the complexity level. *Brain and Behavioral Sciences, 13*(3), 423-469. doi:10.1017/S0140525X00079577.

Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of face perception. *Annual Review of Neuroscience*, 31: 411-437

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*, 161-204. doi:10.1080/14640749108400966

Valentine, T., Darling, S., & Donnelly, M. (2004). Why are average faces attractive? The effect of view and averageness on the attractiveness of female faces. *Psychonomic Bulletin & Review, 11*(3), 482-487. doi:Doi 10.3758/Bf03196599

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of Attention and Emotion on Face Processing in the Human Brain. *Neuron, 30*, 829-841. doi:10.1016/s0896-6273(01)00328-2

Walker, D., & Vul, E. (2014). Hierarchical encoding makes individuals in a group seem more attractive. *Psychol Sci, 25*, 230-235. doi:10.1177/0956797613497969

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E. J., van Doorn, J., Šmíra, M., Epskamp, S., Etz, A., Matzke, D., de Jong, T., van den Bergh, D., Sarafoglou, A., Steingroever, H., Derks, K., Rouder, J. N., … Morey, R. D. (2017). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic bulletin & review, 25*(1), 58-76.

Webster, M. A. (2011). Adaptation and visual coding. *J Vis, 11*. doi:10.1167/11.5.3

Webster, M. A. (2014). Probing the functions of contextual modulation by adapting images rather than observers. *Vision Research, 104*, 68-79. doi:10.1016/j.visres.2014.09.003

Webster, M. A. (2015). Visual Adaptation. *Annual Review of Vision Science, 1*, 547-567. doi:10.1146/annurev-vision-082114-035509

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428*(6982), 557-561. doi:10.1038/nature02420

Webster, M. A., & MacLeod, D. I. A. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*, 1702-1725. doi:10.1098/rstb.2010.0360

Whitney, D., & Levi, D. M. (2011). Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences, 15*, 160-168. doi:10.1016/j.tics.2011.02.005

Whitney, D., & Leib, A. (2017). Ensemble Perception. *Annu Rev Psychol*. doi:10.1146/annurev-psych-010416-044232

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods, 42*(3), 671-684. doi:10.3758/Brm.42.3.671

Willis, J., & Todorov, A. (2006). First Impressions Making Up Your Mind After a 100-Ms Exposure to a Face. *Psychological Science, 17*, 7. doi:10.1111/j.1467-9280.2006.01750.x

Wolfe, B. A., Kosovicheva, A. A., Leib, A. Y., Wood, K., & Whitney, D. (2015). Foveal input is not required for perception of crowd facial expression. *Journal of Vision, 15*, 11. doi:10.1167/15.4.11

Wolfe, J. M. (1984). Short test flashes produce large tilt aftereffects. *Vision Research, 24*, 1959-1964. doi:10.1016/0042-6989(84)90030-0

Wurtz, R. H., Goldberg, M. E., & Robinson, D. L. (1982). Brain Mechanisms of Visual-Attention. *Scientific American, 246*(6), 124-&. doi:DOI 10.1038/scientificamerican0682-124

Xu, H., Dayan, P., Lipkin, R. M., & Qian, N. (2008). Adaptation across the Cortical Hierarchy: Low-Level Curve Adaptation Affects High-Level Facial-Expression Judgments. *Journal of Neuroscience, 28*, 3374-3383. doi:10.1523/jneurosci.0182-08.2008

Yan, X., Andrews, T. J., & Young, A. W. (2016). Cultural similarities and differences in perceiving and recognizing facial expressions of basic emotions. *Journal of Experimental Psychology: Human Perception and Performance, 42*(3), 423.

Yang, H., Shen, J. H., Chen, J. A., & Fang, F. (2011). Face adaptation improves gender discrimination. *Vision Research, 51*(1), 105-110. doi:10.1016/j.visres.2010.10.006

Ying, H., & Xu, H. (2017). Adaptation reveals that facial expression averaging occurs during rapid serial presentation. *Journal of Vision, 17,* 15. doi:10.1167/17.1.15

Young, A. W., & Bruce, V. (2011). Understanding person perception. *British Journal of Psychology, 102*, 959-974. doi:10.1111/j.2044-8295.2011.02045.x

Zhang, M. S., Wang, X. L., & Goldberg, M. E. (2014). A spatially nonselective baseline signal in parietal cortex reflects the probability of a monkey's success on the current trial. *Proceedings of the National Academy of Sciences of the United States of America, 111*(24), 8967-8972. doi:10.1073/pnas.1407540111

Zhao, L., & Chubb, C. (2001). The size-tuning of the face-distortion after-effect. *Vision Research, 41*(23), 2979-2994. doi:Doi 10.1016/S0042-6989(01)00202-4

Zhao, Y., Zhen, Z., Liu, X., Song, Y., & Liu, J. (2017). The neural network for face recognition: Insights from an fMRI study on developmental prosopagnosia. *Neuroimage, 169*, 151-161. doi:10.1016/j.neuroimage.2017.12.023