Ensemble statistics shape face adaptation and the cheerleader effect

Ying, Haojiang; Burns, Edwin; Lin, Xinyi; Xu, Hong

2019

Ying, H., Burns, E., Lin, X., & Xu, H. (2019). Ensemble statistics shape face adaptation and the cheerleader effect. *Journal of Experimental Psychology*, 148(3), 421–436.
doi:10.1037/xge0000564

<https://hdl.handle.net/10356/145719>

<https://doi.org/10.1037/xge0000564>

© American Psychological Association, 2019. This paper is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. Please do not copy or cite without author's permission. The final article is available, upon publication, at: 10.1037/xge0000564

Downloaded on 26 Apr 2023 17:07:28 SGT

Ensemble statistics shape face adaptation and the cheerleader effect

Haojiang Ying¹, Edwin Burns^{1,2}, Xinyi Lin¹, and Hong Xu^{1*}

1. Psychology, School of Social Sciences, Nanyang Technological University, Singapore

2. Department of Psychology, University of Richmond, USA

*Correspondence:

Dr. Hong Xu

48 Nanyang Ave, HSS-04-06

Psychology

School of Social Sciences

Nanyang Technological University

Singapore 639818

Phone: +65 6592-1571

Fax: +65 6795-5797

Email: xuhong@ntu.edu.sg

Word count: 9717.

Abstract word count: 247.

Abstract

When confronted with a scene of emotional faces, our brains automatically average the individual facial expressions together to create the gist of the collective emotion. Here, we tested whether this ensemble averaging could also occur for facial attractiveness, and in turn shape two related face perception phenomena: adaptation and the cheerleader effect. In our first two experiments, we showed that adaptation aftereffects could indeed be shaped by ensemble statistics; viewing an increasingly unattractive group of faces conversely increased attractiveness judgments for a subsequently presented face. Not only did group adaptation aftereffects occur, but their effects were equivalent to those observed from the morphed average face of the group, suggesting that the visual system had averaged the group together. In our last two experiments, we showed that viewing a target face in an increasingly unattractive group led to the target being perceived as increasingly more attractive: a ‘cheerleader’ effect. Moreover, our results suggest that this cheerleader effect likely comprises of both a social positive effect and a contrastive process, requiring variance between the surrounding and target faces; i.e., the visual system appeared incapable of boosting a target’s attractiveness when all of the faces in the scene were identical. Furthermore, the mean group attractiveness ratings strongly predicted both the cheerleader effect and adaptation aftereffects, with the latter two also interrelated. This suggests that ensemble statistics is the common underlying process linking each of these phenomena. In order to be perceived as beautiful, being surrounded by unattractive friends may help.

Keywords: facial attractiveness, cheerleader effect, visual adaptation, ensemble coding, friend effect

Introduction

One of the most important and widely studied facial traits is that of attractiveness due to its social and evolutionary significance (Little, Jones, & DeBruine, 2011; Rhodes, 2006; Thornhill & Gangestad, 1999; Willis & Todorov, 2006). Research has uncovered numerous attributes that modulate a face's attractiveness (Little et al., 2011; Rhodes, 2006), including: symmetry (Perrett, Burt, Penton-Voak, Lee, Rowland, & Edwards, 1999; Rhodes, Yoshikawa, Clark, Lee, McKay, & Akamatsu, 2001; Grammer & Thornhill, 1994), masculinity/femininity (Perrett, Lee, Penton-Voak, & Rowland, 1998; Perrett, May, & Yoshikawa, 1994) and averageness (DeBruine, Jones, Unger, & Little, 2007; Deffenbacher, Vetter, Johanson, & O'Toole, 1998; O'Toole, Price, Vetter, Bartlett, & Blanz, 1999; Perrett et al., 1994; Rhodes & Tremewan, 1996; Rhodes et al., 2001; Grammer & Thornhill, 1994).

However, 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, a visual phenomenon known as the cheerleader effect (Carragher, Lawrence, Thomas, & Nicholls, 2018; Walker & Vul, 2014; but see Ojio et al., 2015). As it was suggested to us that the label 'cheerleader' 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 makes subsequently presented faces appear more attractive, an attractiveness adaptation aftereffect (Rhodes, 2006). There is a question, however, as to whether the mean attractiveness of a group of faces affects the attractiveness judgments of a target face. More specifically, do we look better when we surround ourselves with attractive or unattractive friends? 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 bar, club 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 for 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 (Walker & Vul, 2014), would be a 'basking in reflected glory' or averaging effect. This predicts that faces are more attractive when judged in a crowd due to the crowd biasing the perception of that face towards the group's average. 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, we could have a contrastive effect, similar to what occurs during face adaptation (Calder, Jenkins, Cassel, & Clifford, 2008; Hsu & Young, 2004; Leopold, O'Toole, Vetter, & Blanz, 2001; Pegors, Mattar, Bryan, & Epstein, 2015; Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003; Rhodes & Leopold, 2011; Webster, Kaping, Mizokami, & Duhamel, 2004). In this case, averaging produces a new norm from the crowd's mean which then boosts 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 when in a crowd. This contrastive effect is because the unattractive group is thought to habituate the neurons that are tuned for

105 unattractive faces, **leading to** subsequent faces to be perceived as more attractive (O'Doherty
106 et al., 2003; Pegors et al., 2015; Rhodes et al., 2003; Webster & MacLeod, 2011).

107 While the averaging and contrastive effects would arguably modulate the **friend** effect
108 in line with the group's mean attractiveness, the two effects' relationships with the group's
109 mean attractiveness would be in opposite directions to one another. For example, under the
110 averaging account, increasingly attractive friends would increase the target's perceived
111 attractiveness. This is because in the averaging theory, the target face becomes more
112 attractive as it is biased towards the mean of the crowd when averaged into the group.
113 Support for this suggestion comes from prior work that shows a Gabor patch's orientation
114 will contribute towards the group's average, yet also be perceived like that average (Morgan,
115 Chubb, & Solomon, 2008; Ross & Burr, 2008). Converse to this, the contrastive hypothesis
116 predicts that increasing group attractiveness would make a target face less attractive. Under
117 such circumstances, the target face may or may not be incorporated into the crowd, but it is
118 compared to the group's mean. This contrastive effect is similar to that arising during
119 adaptation (O'Doherty et al., 2003; Pegors et al., 2015; Rhodes et al., 2003).

120 Finally, the mere presence of a face in a group may be sufficient to cause some kind
121 of social positive effect. In this case, simply being in a group makes a target face more
122 attractive irrespective of how the group looks. What might cause this effect is unclear, but it
123 may be driven by some kind of high level process that can attribute popularity to a face due
124 to the simple fact that it is surrounded by other people. However, it is possible that such a
125 social positive effect might co-occur with either an averaging or contrastive effect. If this
126 were to be true, then we might expect the target face in a group to always be rated as more
127 attractive than when viewed in isolation, but the strength of this **friend** effect may vary in
128 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 in order to obtain its gist (Haberman & Whitney, 2007, 2009; Sweeny & Whitney, 2014; Whitney & Leib, 2017; Ying & Xu, 2017). These ensemble statistics or representations have been shown to arise in the processing of facial identity (Haberman, Brady, & Alvarez, 2015; Leib et al., 2012), gender (Haberman & Whitney, 2007), viewpoint (Sweeny & Whitney, 2014), and emotion (Elias, Dyer, & Sweeny, 2017; Haberman & Whitney, 2009; Wolfe, Kosovicheva, Leib, Wood, & Whitney, 2015; Ying & Xu, 2017). It is time consuming and effortful to assess the attractiveness of a target face compared to all of the other faces in a group (Tsotsos, 1990). It is thus reasonable to imagine that such ensemble coding 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). Evolutionary theory would arguably predict this template should arise in order to assist us in picking the best available mate relative to other options currently available in any given environment. Under these circumstances, we should expect increasingly unattractive groups to bias our judgments for finding the target face as more attractive. This is because someone should be judged more desirable in an unattractive crowd as they are 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.

Experiment 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 seem more attractive or unattractive? 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). To our knowledge, our present study will be the first that tests spatially presented ensemble adaptation of faces. 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 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.

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 ensemble coding via 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 $1 - \beta = 1$. This study and all following experiments in this paper were reviewed and approved by the Ethics Committee of the Psychology and Institutional Review Board (IRB) of Nanyang Technological University. All participants gave their informed and written consent prior to the study.

Apparatus

Visual stimuli were presented on a 17-inch Philips CRT monitor (refresh rate 85 Hz, spatial resolution 1024×768 pixels). The monitor was controlled by an iMac Intel Core i3 computer running Matlab R2010a (Mathworks) via Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). Each subject was seated in a dimly lit room with their head rested on a chin rest 75 cm in front of the monitor. Each pixel subtended 0.024° on the screen.

Stimuli

All of the visual stimuli were female faces chosen from the Oslo face database (Chelnokova et al., 2014). We chose this face database for two reasons. First, it contains 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 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. The attractiveness of each face was assessed on a 7-point scale (1 for most unattractive and 7 for most attractive). All faces were presented individually on the screen 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 by 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 to be used for our adapting and test faces based on these ratings. Due to publishing restrictions, we use faces from Karolinska Directed Emotional Faces database (Lundqvist, Flykt, & Öhman, 1998) for demonstration in our figures.

227 *Test Faces*

228 Using MorphMan 2016 (STOIK Imaging, Moscow, Russia) we morphed the top two
229 attractive faces together to create our most attractive test face. We did this as we wanted to
230 make our highly attractive face even more attractive, and face averaging achieves this goal as
231 average faces are rated more attractive than their non-average counterparts (DeBruine et al.,
232 2007; Perrett et al., 1994). This face was then morphed with the most unattractive face to
233 create a sequence of seven, incrementally spaced, morph continua test faces; we did not
234 average the most unattractive faces as we wanted to have the most unattractive face possible.
235 Based on a small pilot test ($n = 5$), we found participants (who did not participate in any of
236 the experiments reported here) preferred this stream of faces rather than a stream of faces
237 created simply from the most attractive face. The morphed faces were separated by units in
238 proportions of $1/7^{\text{th}}$ attractiveness. For example, the most attractive morph which contained
239 100% of the synthesized attractive face (and 0% of the unattractive face) was equal to 1
240 attractiveness unit; the least attractive morph with 0% of the synthesized attractive face and
241 100% of the unattractive face was equal to 0 attractiveness units. All test stimuli subtended a
242 horizontal and vertical visual angle of $1.80^{\circ} \times 2.40^{\circ}$ respectively.

243 *Adapting Faces*

244 The adapting stimuli were based on the six most attractive faces and the six most
245 unattractive faces (excluding the most unattractive face which was used as the most
246 unattractive test face) from the Oslo face database. There are three types of adapting faces: all
247 6 most attractive faces; all 6 most unattractive faces; or a mixture of 3 most attractive and 3
248 most unattractive faces (randomly selected from the 6 attractive and 6 unattractive faces).
249 These faces were displayed at the same size as the test faces.

250

Procedure

We used a block design comprising of 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. On each trial, the test stimulus presented was one of the morphed faces selected at random. Each trial began with a centrally presented fixation cross for 500 ms. This cross would be ever present on 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 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 (much briefer than the adaptors). It has been documented in the face adaptation literature that presenting a test face at a shorter duration could enhance the adaptation aftereffect (e.g., Burton, Jeffery, Bonner, & Rhodes, 2016; Rhodes, Jeffery, Clifford, & Leopold, 2007). 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 between the adapting and test faces on the computer screen, 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 appropriate keyboard response to indicate whether the test face was attractive or unattractive ('A' for attractive, 'S' for unattractive). This screen commenced with a 50 ms

275 beep noise to alert participants to respond, with their response starting the next trial (Figure
 276 2).

277

278 Figure 1. An example of the **adapting** faces used in Experiment 1 (the demonstrated faces are
 279 AF01NES, AF06NES, AF08NES, AF17NES, AF20NES, and AF34NES from KDEF database). (A)
 280 The 6 **adapting** faces formed a hexagon. In the experiment, the central fixation cross was right in the
 281 center of them. (B) The schematic of the relative locations of the stimuli. The **test** face was presented
 282 in the central position (the intersection of the three lines) of Cartesian coordinates. The locations of
 283 the central points of the **adapting** faces are at the end points of each line, the coordinates are the
 284 relevant location for each **adapting** face. For example, the top left **adapting** face is 1.13° to the left and
 285 2.16° to the top of the **fixation cross**.

286

287 Within each block, each **test** face was presented 10 times in a random sequence
 288 making a total of 70 trials in each block. Similarly, the locations of the six **adapting** face
 289 identities were also shuffled randomly around the hexagon. Each block lasted around seven
 290 minutes, and there was a seven-minute rest in between **consecutive** blocks to avoid any
 291 carryover effects. The order of the blocks was randomized for each subject. Data collection
 292 started after subjects had sufficient practice trials to familiarize themselves with the task.

Figure 2. The sequence of an adaptation trial (the demonstrated faces are AF01NES, AF06NES, AF08NES, AF17NES, AF20NES, AF26NES, and AF34NES from KDEF database). Subjects fixated on the cross and pressed the space bar to initiate a trial. After 500 ms, the adaptors, six faces appeared at the screen for 1 s. The locations of the adaptors are the same as the faces in Figure 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 data were sorted into proportions of ‘attractive’ responses to each test stimulus per condition. The test stimuli were parameterized according to the attractiveness units in the morphed test faces. The proportions of ‘attractive’ responses were then plotted against each test face, and the resulting psychometric curves were fitted with a sigmoidal function $f(x) = 1/[1 + e^{-a(x-b)}]$, where $a/4$ is the slope and b gives the test-stimulus parameter corresponding to the 50% point of the psychometric function [the point of subjective equality (PSE)]. The attractiveness aftereffect (or 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 analysis of variance (ANOVA) and pairwise comparisons (with Bonferroni corrections) 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.

We then performed a correlation to examine the extent to which our adaptation aftereffects were being driven by the mean attractiveness of the group of adapting faces. As all participants were repeatedly measured under three adaptation conditions, it is problematic for us to test the relationship between the adaptation aftereffects and the mean ratings of each adaptation condition's faces via a conventional correlational analysis; this is because each participant would provide three data points to each correlation, hence violating the assumption of independence between each observation in the correlation. To solve this problem, we used a repeated measures correlation analysis (Bakdash & Marusich, 2017) to measure the strength of the relationship between adaptation and group attractiveness.

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).

Results and Discussion

The results from all participants averaged together are shown in Figure 3A. We plotted the fraction of 'attractive' responses as a function of the attractiveness units of the test faces. The 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 (blue solid 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).

339

Figure 3. The attractiveness adaptation aftereffects of **adapting** faces with different levels of mean attractiveness (Experiment 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 the 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 aftereffect 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 3B). 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 judged 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 judged 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 (Mauchly's test indicated that the assumption of sphericity was not violated, $\chi^2(2) = .72$, $p = .70$; $F(2,38) = 33.48$, $p < .001$, $\eta_p^2 = .64$). Further Bonferroni corrected pairwise comparisons indicated that there were 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 ($t(19) = 3.16$, $p = .016$, Cohen's $d = .71$). Moreover, 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 3C). A significant association was revealed between these mean attractiveness ratings and the size of the aftereffects ($r = .81, p < .001, 95\% \text{ CI } [0.67, 0.90]$). Ensemble averaging therefore appears to drive the creation of a new attractiveness norm through adaptation.

Experiment 2. Ensemble adaptation aftereffects are equal to their average counterparts

Experiment 1 showed that a group of faces could produce adaptation aftereffects. We wondered if this occurred from an averaging or summation process. In Experiment 2 we tested these possibilities in a number of different ways. First, we generated a morphed average face (Figure 4B) of the attractive face group (Figure 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 facial 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 4C) from the group (Figure 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, adapting to a group of the same single face presented in six locations at the same time (Figure 4D) should result in larger aftereffects than that produced by the single face (Figure 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 1.

Methods

Subjects, Apparatus, Stimuli, and Procedure

Thirty new participants (20 females, mean age 23.93 years) took part in this experiment. We chose this sample size for two reasons. The first reason was that the result of a power analysis (using G*Power 3.1 software, basing on $\eta_p^2 = .64$ from Experiment 1, with α -value at .05, and power $(1 - \beta)$ at .95) indicated that we would need at least 6 participants. We further considered that Pegors and colleagues (2015) used 30 as sample size in their facial attractiveness study. We therefore chose 30 as the sample size for the current study.

We used the same lab setting, procedure, analysis, and the face dataset as in Experiment 1, except for a couple of changes. First, we wanted to rule out any possible confounding influence of the fact that the test faces and adapting faces were derived from the same identities. We therefore created a new test face stream and collection of adapting faces. In this experiment, the new test faces comprised the morph continua of the most attractive face and the least attractive face from the previous larger face database (based upon the judgements from Experiment 1). The adapting faces were the six most attractive faces taken from the remaining group of the database after the test face had been removed. Therefore, the testing faces and the adapting faces were from different identities. As in Experiment 1, the faces were cropped with an oval shape mask. Moreover, the luminance of the faces was further equalized by the SHINE toolbox (Willenbockel et al., 2010).

Secondly, we have four different adaptation conditions (Figure 4). For the average attractive condition (Figure 4B, AVE), 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 4A) together. For the Single1 condition, we picked one of the faces 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 the 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 4D).

Figure 4. The adapting faces for Experiment 2 (the 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 scramble faces (Single1) condition. (D) The single face repeated six-time (Single6) condition.

Results and Discussion

The results from all participants averaged together are shown in Figure 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.

Figure 5. The attractiveness adaptation aftereffects produced by each of the four conditions in Experiment 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 bars indicate 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 during adaptation (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. 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 the two-tailed paired t tests. Note that, a more positive adaptation aftereffect measured by PSE shift indicates that the test faces were perceived as less attractive than on their own.

The summary of the adaptation aftereffects measured by PSE shift is illustrated in Figure 5B. 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 generated 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 ($\epsilon = .73$);

$F(2.18, 63.13) = 9.72$, $p < .001$, $\eta_p^2 = .25$). 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 ($t(29) = .124$, $p = 1$, Cohen's $d = .023$). Moreover, both ATT and AVE conditions generated significantly larger aftereffects than both Single1 and Single6 conditions (all $ps < .031$). 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 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 1/6th of the ATT group's aftereffects, we performed a one-sample t -test on the Single1 condition's aftereffects, comparing to the 1/6th 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 Experiments 1 and 2 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. For example, adapting to the morphed average face (AVE) from the attractive group led to equivalent aftereffects to those produced by the attractive group (ATT). 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?

Experiment 3. We look better with unattractive friends

In our previous experiments in this paper, we found that ensemble perception 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 (Carragher et al., 2018; 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 emotion adaptation studies in an RSVP sequence (Ying & Xu 2017). We therefore 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, we used the same faces from our first experiment (Experiment 1) and our present experiment (Experiment 3), so that we could directly test 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. 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, Jones, Unger, & Little, 2007; Perrett, May, & Yoshikawa, 1994). Alternatively, the friend effect might be explained by a contrastive adaptation effect as we saw in Experiment 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 both 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 an individual's attractiveness irrespective of their looks. This latter effect may, however, also occur concurrently with the averaging or contrastive theories we described above. We tested these hypotheses in Experiment 3.

Methods

Subjects, Apparatus, Stimuli, and Procedure

The same subjects ($n = 20$) from Experiment 1 participated in Experiment 3. From a post-hoc power analysis (with α -value of .05, $\eta_p^2 = .35$, G*Power 3.1), we found that this sample size was sufficient to yield high power: $1 - \beta = .98$.

We used the same lab setting and stimuli as in Experiment 1, except for a couple of minor adjustments to the paradigm. The key difference was that in the present paradigm, the surrounding faces were presented at the same time as the target face (1 second, Figure 6). The spatial arrangement of the surrounding faces was the same as in Experiment 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 would 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 6. The sequence of one example trial (the demonstrated faces are AF01NES, AF06NES, AF08NES, AF17NES, AF20NES, AF26NES, and AF34NES from KDEF database). Subjects fixated on the cross and pressed the space bar to initiate a block. 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 parameters for all conditions and experiments are detailed in the Methods section.

Results and Discussion

The results from all participants averaged together are shown in Figure 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 (blue solid 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). Moreover, it appeared that the decreasing 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 7B. Compared to the baseline PSE, the unattractive ($M = 12.70\%$, $SEM = .027$; $t(19) = 4.64$, $p < .001$, Cohen’s $d = 1.04$, 95% CI [0.07, 0.18]), mixed ($M = 9.21\%$, $SEM = .025$; $t(19) = 3.71$, $p = .002$, Cohen’s $d = 0.83$, 95% CI [0.04, 0.14]), 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 repeated measures analysis of variance (ANOVA) also indicated significant differences among all three friend conditions (Mauchly’s test indicated that the assumption of sphericity was not violated, $\chi^2(2) = 2.21$, $p = .33$; $F(2,38) = 10.34$, $p < .001$, $\eta_p^2 = .35$). Further Bonferroni corrected pairwise comparisons indicated that the unattractive condition produced the largest friend effect: larger than the attractive ($t(19) = 3.93$, $p = .003$, Cohen’s $d = 0.88$) and marginally larger than mixed friend ($t(19) = 2.33$, $p = .086$, Cohen’s $d = 0.53$) conditions. Similarly, greater friend effects were found in the mixed over the attractive condition ($t(19) = 2.62$, $p = .050$, Cohen’s $d = 0.59$).

Figure 7. The effects of surrounding faces with different levels of group attractiveness (Experiment 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 indicate standard error of the mean. '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 shown for each condition in the figure was calculated using two-tailed paired t tests. 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 for each individual subject. Each data point is derived from the mean attractiveness rating of the surrounding faces and their friend effect from a single observer for each condition. Thus, each participant has his/her own correlation line fitted to the data points, of the same color. Triangles represented the individual subjects' mean ratings of attractive surrounding faces. Squares for the mixed surrounding, and circles for the unattractive surrounding. Taken together, the size of friend effect and the mean

attractiveness rating correlated significantly by repeated measure correlation ($r = .63, p < .001, 95\%$ CI [0.39, 0.79]). (D) The attractiveness aftereffect from each condition (relative to the no adaptation baseline condition) plotted as a function of the corresponding condition's friend effect (calculated relative to the no friends baseline). Adaptation aftereffects and the friend effects were significantly correlated by repeated measure correlation ($r = .65, p < .001, 95\%$ CI [0.42, 0.80]).

We specifically predicted that the friend effect may be influenced by ensemble perception by forming a new norm from the group's mean attractiveness ratings. To support this hypothesis (Figure 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 1) and contextual (Experiment 3) effects created via **hypothesized** ensemble coding, a further repeated measures correlation (Figure 7D) between the adaptation aftereffects in Experiment 1, and the friend effects from Experiment 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 **what we believe is** 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 8). Both adaptation aftereffects and friend effects were significantly correlated with the mean attractiveness of the face groups in the surround ($r = .81, p < .001$; and $r = .63, p < .001$, respectively). This suggests that the mean attractiveness of a face group is the common factor **associated with** both friend effects and attractiveness adaptation aftereffects.

Figure 8. The generalized model from our two experiments. The mean attractiveness of the face group (ensemble representation) could predict the adaptation aftereffect (Experiment 1) and the friend effect (Experiment 3). Moreover, the adaptation aftereffect and the friend effect, which were both associated with ensemble representations, correlated with one another.

In summary, the results from Experiment 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). Why Ojiro and colleagues did not find significant effects is unclear to us, however, our own work and that of others (Carragher et al., 2018; Walker & Vul, 2014) provide, in our opinion, strong evidence that the friend effect is a true effect. Moreover, we found on average across the three experimental conditions a large effect size relative to the no 'friends' baseline (mean Cohen's $d = 0.82$), compared to the effect size in the other studies (Walker & Vul (2014, Exp 4): $\eta_p^2 = .197$; Carragher et al., (2018): mean Cohen's $d = .56$; Ojiro et al., (2015, Exp 2): $\eta_p^2 = .017$). The large effect size in our study was likely due to our employment of a psychophysical measure of behavioral performance that is highly sensitive at detecting perceptual effects (PSE).

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 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 a 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.**

Experiment 4: The friend effect arises as a contrastive effect when ensemble perception is engaged

In Experiment 4, 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 the contrast between the target and ensemble perception of the friends, then we might fail to find a friend effect when all faces are identical. Hints for this possibility come from our adaptation results in Experiment 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 processes as adaptation, then we may not find a friend effect when all faces are identical. This is because we would expect the same neuronal 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 1 and 2, as well as test faces from Experiment 2, in order to replicate the friend effects we found in Experiment 3. Finally, we used different facial identities between adaptation and test in order to remove any possible confound of identity, much like how we changed the stimuli for Experiment 2 in order to counter the same issue in Experiment 1.

Methods

Subjects, Apparatus, Stimuli, and Procedure

Thirty new subjects (22 females, mean age 25.03 years) participated in this experiment. We chose 30 as the sample size for several reasons. First, using the effect size of Experiment 3 ($\eta_p^2 = .35$), a power analysis (using G*Power 3.1, with α -value at .05, and power ($1 - \beta$) at .95) indicated that we needed at least 13 participants. Considering the sample sizes of attractiveness studies using the similar paradigms ($n_{mean} = 27.8$ in Walker & Vul, 2014, ranging from 18-39; $n = 30$ in Pegors et al., 2015), we believed that 30 participants would be sufficient for the experiment.

We used the same stimuli as in Experiment 2, albeit with the unattractive adapting faces from Experiment 1 for our current unattractive and mixed conditions. The paradigm was also largely the same as Experiment 3, except we added a condition where the 6 surrounding faces were identical to the target face. Thus, in the new ‘Same Surrounding’ condition, the target face was always surrounded by six copies of itself.

Results and Discussion

The results from all participants averaged together are shown in Figure 9A. We replicated the findings from Experiment 3 in that the target faces were perceived as more attractive than on their own (baseline with no surrounding faces, black dash-dotted line) when surrounded by the unattractive faces (blue solid line), mixed faces (cyan dashed line), or attractive faces (magenta dotted line). However, the ‘Same Surrounding’ condition (green solid line) showed little shift from the baseline condition, which means that presenting identical friends and target faces in a scene fails to elicit a friend effect.

The summary of the friend effect measured by PSE shift is illustrated in Figure 9B. 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 significantly boosted the attractiveness of the target face. 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 [-0.005, 0.03]). A repeated measures analysis of variance (ANOVA) 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 ($\epsilon = .74$); $F(2.23, 64.66) = 35.12$, $p < .001$, $\eta_p^2 = .55$). Further Bonferroni corrected pairwise comparisons indicated that the unattractive condition produced the largest friend effect: larger than the mixed ($t(29) = 4.04$, $p = .002$, Cohen's $d = .74$), attractive surrounding ($t(29) = 6.02$, $p < .001$, Cohen's $d = 1.10$), and the same surrounding condition ($t(29) = 7.72$, $p < .001$, Cohen's $d = 1.41$). Similarly, greater friend effects were found in the mixed over the attractive condition ($t(29) = 4.02$, $p = .002$, Cohen's $d = .73$) and the same surrounding condition ($t(29) = 6.09$, $p < .001$, Cohen's $d = 1.11$). Finally, the attractive condition also exhibited larger effects when compared to the same surrounding condition ($t(29) = 3.50$, $p = .009$, Cohen's $d = .64$).

We further validated that the friend effect was influenced by ensemble perception by forming a new norm from the surrounding group's mean attractiveness ratings. Using a repeated measures correlation analysis, we found a significant positive correlation (Figure 9C) between the mean attractiveness rating for each surrounding group and the friend effect ($r = .54$, $p < .001$, 95% CI [0.33, 0.70]).

In summary, this study replicated the findings of Experiment 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.

Figure 9. The effects of surrounding faces with different levels of group attractiveness (Experiment 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 indicate standard error of the mean. '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 values shown above each condition in the figure was calculated using two-tailed paired t tests. **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 by repeated measures correlation. Triangles represented the individual subject's mean ratings of attractive surrounding faces and the friend effect. Squares for the mixed surrounding, and circles for the unattractive surrounding. Each color indicates one participant. Taken together, the size of friend effect and the mean attractiveness rating correlated significantly, as in Experiment 3.

General Discussion

In our first two experiments, we showed that ensemble statistics of previously viewed groups could shape subsequent attractiveness judgments **through** 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 Experiments 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 an individual being more likely judged 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 **against** which target faces can be implicitly judged. 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 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 our experiments, subjects were presented **with** natural faces without distortion, yet we still observed large adaptation aftereffects (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 in a group. 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 Experiments 1 and 2. Such incongruence between adaptors and the test face is typically thought to counteract low level retinotopic effects (Adams, Gray, Garner, & Graf, 2010; Leopold et al., 2001) and thus indicates 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 2 we further confirmed that it is the ensemble coding of the crowd that drives adaptation aftereffects. The individual face alone generated much smaller adaptation aftereffects than the face group or their morphed average. 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 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 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 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 4 not only replicated the major findings from Experiment 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 as the attractiveness of the surrounding group increases? 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 1 & 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 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 lack of a friend effect in the identical face (the ‘Same Surrounding’ condition) in Experiment 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 2. Based upon our current experiments, however, we are unable to answer whether or not the target face is included in the ensemble representation.

Fischer and Whitney (2011) showed their stimuli in a similar hexagon fashion as Experiment 2 but in a crowding paradigm. They found that in their experiment, the emotion of the central face was averaged towards the surrounding faces. By contrast, we found that the attractiveness of the central face was modulated by the ensemble representation of the surrounding faces in the opposite direction. We believe this discrepancy stems from one key difference between the experiments. Fischer and Whitney (2011) *positioned their face sets at eccentricities known to promote visual crowding* (Whitney & Levi, 2011); we, however, put the set of faces in the center of the screen close to the target face. This may suggest that ensemble perception favors extra-foveal information more (Wolfe et al., 2015), *and the*

866 receptive field of neurons increases with eccentricity (Kay et al., 2015). There is a consensus
867 that ensemble perception compresses the noisy and redundant information in the view
868 (Alvarez, 2011). Our Experiment 3 complements such a notion: the task related foveal input
869 is not redundant, but survives (at least partially) from the ensemble representation of task
870 irrelevant extra-foveal information.

871 It is remarkable that ensemble perception occurs in adaptation and the friend effect
872 experiments even though participants were never instructed to look directly at the group of
873 surround faces, and in the case of Experiments 3 and 4, only paid attention to the single target
874 face in the center of the face crowd. This suggests that ensemble perception can occur
875 regardless of directed attention and supports the claim that such perception is an involuntary
876 process. Similarly, the fact that our participants never look directly at the faces would seem to
877 indicate that they were not simply ‘picking’ a single face out of the group to base their
878 judgments upon. This suggestion is further supported by the correlations between the mean
879 attractiveness ratings of the underlying groups and the friend/adaptation effects. Our current
880 study supports the suggestion that ensemble perception occurs for facial attractiveness
881 adaptation, in addition to emotion adaptation (Ying & Xu, 2017). These findings fit with the
882 view that ensemble perception likely arose through an evolutionary advantage at being able
883 to judge the attractiveness of any particular face against recent and current experiences.
884 However, such high-level ensemble coding is likely to have arisen from a precursor system
885 that initially processes lower level information such as contour and textures in the
886 environment. One could imagine that as man became a more social creature, then this process
887 was developed for complex, higher level social information such as attractiveness. In any
888 case, this perspective suggests that other aspects of facial social traits (Oosterhof & Todorov,
889 2008), such as trustworthiness or dominance, might also be susceptible to ensemble encoding.

In summary, we have provided evidence that supports the ubiquitousness of ensemble coding in shaping face perception. We have 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.

Context of the Research

Our lab has been studying face perception since it was founded in 2010. We became interested in facial attractiveness adaptation from 2014 as the university provided funding to support top undergraduate students to gain research experience in different labs through the Undergraduate Research Experience on Campus (URECA) scheme. Research projects on facial attractiveness were popular for students (URECA 2014 – 2018) and gradually led to increasing evidence on the nature of facial attractiveness adaptation (Tan & Xu, [URECA 2014-15](#), APCV 2015; Kan, Ying, & Xu, URECA 2015-16). This prior work led to revised project designs (Ying et al., VSS 2017) before resulting in the studies found in the present manuscript. All of the authors have a general interest in understanding how face perception works. The role that ensemble statistics plays in shaping face perception is still very much open to debate. We were therefore curious if it could shape two phenomena related to face perception: the adaptation and friend effect. If ensemble statistics do shape facial attractiveness judgments, then this would have practical implications in how we may want to appear say through a dating app. For example, are we more likely to be appraised as being attractive when we are sharing a photograph with good looking or unattractive friends? As

we expected, both the friend effect and face adaptation were influenced by the underlying group attractiveness, whereby unattractive ‘friends’ appeared to result in others finding us more attractive. If an individual wants to look beautiful, they should surround themselves with unattractive friends.

Author Contributions

H. Ying, X. Lin, E. Burns, and H. Xu developed the study concept and contributed to the study design. Testing and data collection were performed by H. Ying and X. Lin. H. Ying performed the data analyses and interpretation under the supervision of H. Xu. H. Ying drafted the manuscript, and E. Burns and H. Xu provided critical revisions. All authors approved the final version of the manuscript for submission.

Acknowledgements

Supported by a Nanyang Technological University Research Scholarship (HY), a College of Arts, Humanities and Social Sciences Incentive Scheme (HX), and a Singapore Ministry of Education Academic Research Fund (AcRF) Tier 1 (HX). Concept of Exp 1 was presented at Asian-Pacific Conference on Vision (APCV), Aug 2015, Singapore. Parts of this research (data from Exp 1 & 3) were presented at the Annual Meeting of Visual Science Society (VSS), May 2017, St. Pete Beach, Florida. The research reported here forms part of H. Ying’s PhD thesis at Nanyang Technological University. All data have been made publicly available via the Open Science Framework (OSF) and can be accessed at <https://osf.io/9q7uh/>. We thank Ms. Nadine Garland for proofreading the manuscript.

Reference

- Abbas, Z. A., & Duchaine, B. (2008). The role of holistic processing in judgments of facial attractiveness. *Perception*, 37(8), 1187-1196.
- Adams, W. J., Gray, K. L., Garner, M., & Graf, E. W. (2010). High-level face adaptation without awareness. *Psychol Sci*, 21(2), 205-210. doi:10.1177/0956797609359508
- Afraz, S. R., & Cavanagh, P. (2008). Retinotopy of the face aftereffect. *Vision Res*, 48(1), 42-54. doi:10.1016/j.visres.2007.10.028
- Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends Cogn Sci*, 15(3), 122-131. doi:10.1016/j.tics.2011.01.003
- Anderson, N. H., Lindner, R., & Lopes, L. L. (1973). Integration theory applied to judgments of group attractiveness. *Journal of Personality and Social Psychology*, 26(3), 400-408. doi: 10.1037/h0034441
- Anzures, G., & Mondloch, C. J. (2009). Face Adaptation and Attractiveness Aftereffects in 8-Year-Olds and Adults *Child Development*, 80(1), 14.
- Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in psychology*, 8, 456. doi: 10.3389/fpsyg.2017.00456.
- Barracough, 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.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat Vis*, 10(4), 433-436.
- 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(3), 921.
- Burns, E. J., Martin, J., Chan, A. H., & Xu, H. (2017). Impaired processing of facial happiness, with or without awareness, in developmental prosopagnosia. *Neuropsychologia*, 102, 217-228.
- Burton, N., Jeffery, L., Bonner, J., & Rhodes, G. (2016). The timecourse of expression aftereffects. *Journal of vision*, 16(15), 1-1.
- Calder, A. J., Jenkins, R., Cassel, A., & Clifford, C. W. (2008). Visual representation of eye gaze is coded by a nonopponent multichannel system. *Journal of Experimental Psychology: General*, 137(2), 244-261.
- Carragher, D. J., Lawrence, B. J., Thomas, N. A., & Nicholls, M. E. (2018). Visuospatial asymmetries do not modulate the cheerleader effect. *Scientific reports*, 8(1), 2548.
- 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. *Mol Psychiatry*, 19(7), 746-747. doi:10.1038/mp.2014.1
- 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
- Cunningham, M. R., Roberts, A. R., Barbee, A. P., Druen, P. B., & Wu, C. H. (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(2), 261.
- DeBruine, L., & Tiddeman, B. (2017). WebMorph., Retrieved from <http://webmorph.org>.
- DeBruine, L., Jones, B. C., Unger, L., & Little, A. C. (2007). Dissociating averageness and attractiveness: attractive faces are not always average. *Journal of Experimental Psychology: Human Perception and Performance*, 33(6), 11. doi:0.1037/0096-1523.33.6.1420.

- Deffenbacher, K. A., Vetter, T., Johanson, J., & O'Toole, A. J. (1998). Facial aging, attractiveness, and distinctiveness. *Perception*, 27(10), 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.
- Elias, E., Dyer, M., & Sweeny, T. D. (2017). Ensemble Perception of Dynamic Emotional Groups. *Psychol Sci*, 28(2), 193-203. doi:10.1177/0956797616678188
- 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.
- Fischer, J., & Whitney, D. (2011). Object-level visual information gets through the bottleneck of crowding. *J Neurophysiol*, 106(3), 1389-1398. doi:10.1152/jn.00904.2010
- Furl, N. (2016). Facial-attractiveness choices are predicted by divisive normalization. *Psychological science*, 27(10), 1379-1387.
- Grammer, K., & Thornhill, R. (1994). Human (*Homo sapiens*) facial attractiveness and sexual selection: the role of symmetry and averageness. *Journal of comparative psychology*, 108(3), 233.
- 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. *J Exp Psychol Gen*, 144(2), 432-446. doi:10.1037/xge0000053
- 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. *J Exp Psychol Hum Percept Perform*, 35(3), 718-734. doi:10.1037/a0013899
- 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.
- Haxby, J. V., & Gobbini, M. I. (2011). Distributed neural systems for face perception. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 93–110). New York: Oxford University Press.
- 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(2), 199.
- Hsu, S. M., & Young, A. (2004). Adaptation effects in facial expression recognition. *Visual Cognition*, 11(7), 871–899.
- 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 of London B: Biological Sciences*, 274(1611), 899-903.
- Leib, A. Y., Puri, A. M., Fischer, J., Bentin, S., Whitney, D., & Robertson, L. (2012). Crowd perception in prosopagnosia. *Neuropsychologia*, 50(7), 1698-1707. doi:10.1016/j.neuropsychologia.2012.03.026
- Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nat Neurosci*, 4(1), 89-94. doi:10.1038/82947

- 1037 Little, A. C., Burt, D. M., Penton-Voak, I. S., & Perrett, D. I. (2001). Self-perceived
1038 attractiveness influences human female preferences for sexual dimorphism and
1039 symmetry in male faces. *Proceedings of the Royal Society of London B: Biological*
1040 *Sciences*, 268(1462), 39-44.
- 1041 Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). Facial attractiveness: evolutionary
1042 based research. *Philos Trans R Soc Lond B Biol Sci*, 366(1571), 1638-1659.
1043 doi:10.1098/rstb.2010.0404
- 1044 Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of
1045 spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal*
1046 *of neurophysiology*, 77(1), 24-42.
- 1047 Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska directed emotional faces
1048 (KDEF). *CD ROM from Department of Clinical Neuroscience, Psychology section,*
1049 *Karolinska Institutet*, 91-630.
- 1050 Luo, A. X., & Zhou, G. (2018). Ensemble perception of facial attractiveness. *Journal of*
1051 *vision*, 18(8), 7-7.
- 1052 Luo, C., Burns, E., & Xu, H. (2017). Association between autistic traits and emotion
1053 adaptation to partially occluded faces. *Vision research*, 133, 21-36.
- 1054 Kay, K. N., Weiner, K. S., Grill-Spector, G. (2015), Attention reduces spatial uncertainty in
1055 human ventral temporal cortex. *Current Biology*, 25 (5), 595-600
- 1056 Morgan, M., Chubb, C., & Solomon, J. A. (2008). A ‘dipper’ function for texture
1057 discrimination based on orientation variance. *Journal of Vision*, 8(11), 9-9.
- 1058 Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings*
1059 *of the National Academy of Sciences*, 105(32), 11087-11092.
- 1060 Ojio, Y., Gobara, A., Nam, G., Sasaki, K., Kishimoto, R., Yamada, Y., & Miura, K. (2015).
1061 Two replications of “Hierarchical encoding makes individuals in a group seem more
1062 attractive (2014; Experiment 4)”. *The Quantitative Methods for Psychology*, 11, r8-
1063 r11.
- 1064 O’Toole, A. J., Price, T., Vetter, T., Bartlett, J. C., & Blanz, V. (1999). 3D shape and 2D
1065 surface textures of human faces: the role of “averages” in attractiveness and age.
1066 *Image and Vision Computing*, 18(1), 9-19.
- 1067 O’Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M., & Dolan, R. J. (2003).
1068 Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness.
1069 *Neuropsychologia*, 41(2), 147-155. doi:10.1016/s0028-3932(02)00145-8
- 1070 Pegors, T. K., Mattar, M. G., Bryan, P. B., & Epstein, R. A. (2015). Simultaneous perceptual
1071 and response biases on sequential face attractiveness judgments. *Journal of*
1072 *Experimental Psychology: General*, 144(3), 664.
- 1073 Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming
1074 numbers into movies. *Spatial vision*, 10(4), 437-442.
- 1075 Perrett, D., Burt, D. M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., & Edwards, R.
1076 (1999). Symmetry and human facial attractiveness. *Evolution and human behavior*,
1077 20(5), 295-307.
- 1078 Perrett, D., Lee, K. J., Penton-Voak, I. S., & Rowland, D. A. (1998). Effects of sexual
1079 dimorphism on facial attractiveness. *Nature*, 394(6696), 884.
- 1080 Perrett, D., May, K. A., & Yoshikawa, S. (1994). Facial shape and judgements of female
1081 attractiveness. *Nature*, 368(6468), 239-242. doi:10.1038/368239a0
- 1082 R Core Team (2017). *R: A language and environment for statistical computing*. R
1083 *Foundation for Statistical Computing*, Vienna, Austria. <https://www.R-project.org/>.
- 1084 Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annu Rev Psychol*, 57,
1085 199-226. doi:10.1146/annurev.psych.57.102904.190208

- 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(6), 558-566.
- Rhodes, G., Jeffery, L., Clifford, C. W., & Leopold, D. A. (2007). The timecourse of higher-level face aftereffects. *Vision research*, 47(17), 2291-2296.
- 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(5), 611-625.
- Rhodes, G., & Leopold, D. A. (2011). Adaptive norm-based coding of face identity. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 263-286). New York: Oxford University Press.
- Rhodes, G., & Tremewan, T. (1996). Averageness, exaggeration, and facial attractiveness. *Psychological science*, 7(2), 105-110.
- Ross, J., & Burr, D. (2008). The knowing visual self. *Trends in Cognitive Sciences*, 12(10), 363-364.
- Stormer, V. S., & Alvarez, G. A. (2016). Attention Alters Perceived Attractiveness. *Psychol Sci*, 27(4), 563-571. doi:10.1177/0956797616630964
- Sweeny, T. D., & Whitney, D. (2014). Perceiving crowd attention: ensemble perception of a crowd's gaze. *Psychol Sci*, 25(10), 1903-1913. doi:10.1177/0956797614544510
- 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
- Thornhill, R., & Gangestad, S. W. (1999). Facial attractiveness. *Trends in cognitive sciences*, 3(12), 452-460.
- Walker, D., & Vul, E. (2014). Hierarchical encoding makes individuals in a group seem more attractive. *Psychol Sci*, 25(1), 230-235. doi:10.1177/0956797613497969
- 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. (2011). Visual adaptation and face perception. *Philos Trans R Soc Lond B Biol Sci*, 366(1571), 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 Cogn Sci*, 15(4), 160-168. doi:10.1016/j.tics.2011.02.005
- Whitney, D., & Yamanashi 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. *Psychol Sci*, 17(7), 592-598. 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. *J Vis*, 15(4), 11. doi:10.1167/15.4.11
- 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(13), 3374-3383.
- Ying, H., & Xu, H. (2017). Adaptation reveals that facial expression averaging occurs during rapid serial presentation. *J Vis*, 17(1), 15. doi:10.1167/17.1.15