Predicting Attractiveness from Face Parts Reveals Multiple Covarying Cues

Chang Hong Liu

Andrew W. Young

J Li

X Tian

Wenfeng Chen

Abstract

In most studies of facial attractiveness perception, judgments are based on whole face images. Here we investigated how attractiveness judgments from parts of faces compare to perceived attractiveness of the whole face, and to each other. We manipulated the extent and regions of occlusion, where either the left/right or the top/bottom half of the face was occluded. We also further segmented the face into relatively small horizontal regions involving the forehead, eyes, nose, or mouth. The results demonstrated the correlated nature of face regions, such that an attractiveness judgment for one face part can be highly predictive of the attractiveness of the whole face or the other parts. The left/right half of the face created more accurate predictions than the top/bottom half. Judgments involving a larger area of the face (i.e., left/right or top/bottom halves) produced more accurate predictions than those derived from smaller regions, such as the eyes or the mouth alone, but even the smallest and most featureless region investigated (the forehead) provided useful information. The correlated nature of the attractiveness of face parts shows that perceived attractiveness is determined by multiple covarying cues that the visual system can exploit to determine attractiveness from a single glance.

Keywords: Facial attractiveness; occlusion; parts; whole; inferences about attractiveness

Public Significance Statement: This study investigates the contribution of different face regions to the perceived attractiveness of a face by asking whether different regions make independent contributions to a face's attractiveness, or whether they offer similar information. Although it is commonly assumed that judging facial attractiveness requires inspection of a full face image, our experiments show that even a relatively small part of a face can be sufficient for an observer to infer how attractive the face is; the level of attractiveness of each part of a face is often highly similar to that of other parts of the face. This facilitates the rapid perception of facial attractiveness in what can sometimes be a relatively chaotic natural environment.

Predicting Attractiveness from Face Parts Reveals Multiple Covarying Cues

Facial attractiveness can be determined from remarkably brief presentations that are effectively a single glance (Bar et al., 2006; Liu & Chen, 2018; Locher et al., 1993; Saegusa & Watanabe, 2016; South Palomares & Young, 2018; Willis & Todorov, 2006). Rightly or wrongly, perceived attractiveness can dominate first impressions and exert a profound influence on our judgements of unfamiliar individuals (Dion et al., 1972; South Palomares & Young, 2019; Todorov et al., 2015).

Research on the perceptual determinants of judgments of facial attractiveness has mainly been conducted within an evolutionary framework and sought to identify specific cues that make a face attractive (Rhodes, 2006). The most well-known of such cues are symmetry (Scheib et al., 1999; Thornhill & Gangestad, 1993), averageness (Langlois et al., 1994; Baudouin & Tiberghien, 2004; Rubenstein et al., 2002; Valentine et al., 2004) and sexual dimorphism (Perrett et al., 1998). The dominant methods used in such studies focus on the perceived attractiveness of whole face images. As such, they seem to involve an implicit assumption that these cues are accrued from the whole face. For example, the role of symmetry is usually studied by comparing attractiveness judgments of original whole faces that are not perfectly symmetrical with the versions of the same faces that are made mirror symmetric through graphic manipulations. The effect of averageness is usually studied by comparing the attractiveness of original whole faces with the average of these faces created through morphing. Similarly, the role of sexual dimorphism is studied by comparing original photographs of whole faces with feminised or a masculinised versions of these faces.

This reliance on whole face stimuli leaves open questions concerning the contributions of different regions of the face and different facial features to overall perceived attractiveness. Whilst a minority of previous studies have investigated the attractiveness of different parts of the face and how these may contribute to initial stages of impression formation (Liu & Chen, 2018; Saegusa & Watanabe, 2016; Orghian & Hidalgo, 2020), they have mainly been concerned with the contribution of different face regions; for example, with whether the eyes have a particularly important role. As such they have not directly addressed the issue of how closely evaluation of the whole face relates to the attractiveness of its parts. Could part of a face itself be sufficient for judging facial attractiveness, or can attractiveness only be inferred from combining information across different face regions? To address this issue, we investigated how well attractiveness judgments from parts of faces can predict perceived attractiveness of the whole face itself, and how far the attractiveness of different parts of faces correlate with each other. Our particular focus was of interest was thus on whether the attractiveness of each part of the face will covary with the other parts, in which case the attractiveness of the whole face should be largely predictable from its constituent parts. Covariation between the attractiveness of different face parts is of theoretical importance because it could be used to support flexible ways of determining attractiveness in what can sometimes be a relatively chaotic natural environment. Moreover, covariation would be useful if the brain relies on predictive coding mechanisms to support face perception (Huang & Rao, 2011; Johnston et al., 2016, 2017). We therefore tested explicitly the extent to which such covariation exists.

Recent studies certainly suggest that multiple cues combine to determine the overall perceived attractiveness of a facial image (Holzleitner et al., 2019; Said & Todorov, 2011; Sutherland et al., 2013; Vernon et al., 2014). However, all of these studies follow the normal approach of presenting whole face images. Here, we introduce a complementary procedure that involves presenting face parts and establishing how closely the attractiveness of the separate parts relates to that of other parts and to the perceived attractiveness of the whole face. This is of interest not only because in everyday lives we sometimes make judgments of attractiveness from partially occluded faces (Liu & Chen, 2018; Orghian & Hidalgo, 2020) - in the words of the song, we may see a stranger across a crowded room - but especially because it can offer a direct approach to addressing questions concerning the role of different cues and any corresponding cue complementarity.

In the case of symmetry, for example, it might be expected that some parts of a face can inform about the whole face simply because the information is roughly duplicated bilaterally. This has been noted in the literature concerning both face identity (Troje & Bülthoff, 1998) and attractiveness (Thornhill & Gangestad, 1993). Scheib et al. (1999) have demonstrated a positive relationship between women's attractiveness ratings of male faces and symmetry, even when only the left or right half of the male faces that contained minimal symmetry cues were rated. However, because their main interest was in how attractiveness ratings of the left or right halves correlated with overall symmetry of the same faces, they did not investigate how closely attractiveness judgments made from incomplete stimuli related to judgements of the whole face. More importantly, little is known about whether perceived attractiveness of a part predicts perceived attractiveness of the other parts or the whole face when information about bilateral symmetry is not sufficient. For instance, although the appearance of one eye may be used to infer the other with the assumption that they have approximately the same shape, information about an eye is not inherently useful for inferring the shape of the nose or the mouth. Inferences made of a partially occluded face or an isolated face part in a case like this are likely to rely on information other than bilateral symmetry.

Since the face can be occluded in many ways in which an exposed region is not a mirror reflection of the hidden part, these offer an interesting test of how far attractiveness can be predicted without invoking bilateral symmetry. Indeed, it seems likely that information other than bilateral symmetry allows for reasonable inferences. For example, the closeness of a face to a population average, which is often regarded as a strong criterion of facial attractiveness (Langlois et al., 1994; Baudouin & Tiberghien, 2004; Rubenstein et al., 2002; Valentine et al., 2004), may be correlated locally in the sense that if one part of a face resembles an average, the other parts are likely to be the same. Viewers may use this as an assumption in their inferential process when they judge attractiveness from an occluded face. Sexual dimorphism, which also contributes significantly to attractiveness (Perrett et al., 1998), may also be inferred from local features of a face in similar fashion. These assumptions about the correlated nature of face parts may be built on prior experience.

The brain may also employ individual properties of a face to infer whole face attractiveness from an exposed region. For example, skin texture and contrast, both being important elements for determining facial attractiveness (Jones et al., 2004; Fink et al., 2006;

Russell et al., 2016), will usually be correlated across different parts of a face. A viewer may therefore use these as a cue for assessing the rest of a face when only a part of it is visible.

The current study, therefore, focused on how observers can use various facial regions to predict whole face attractiveness when only a part of it is exposed, and the extent to which the attractiveness of different parts of the face are correlated with each other. To investigate these issues Experiment 1 assessed inferences of facial attractiveness based on the left/right half of the face, Experiment 2 studied inferences based on top/bottom half of the face, and Experiment 3 investigated these based on isolated facial features, which exposed much smaller areas of the face. Finally, Experiment 4 replicated the main pattern of findings with a different set of stimuli.

Experiment 1

The aim of this experiment was to assess how well attractiveness of the whole face and of the complementary half face could be predicted from the left or right half of the face. This experiment used a similar paradigm to Scheib et al. (1999), but focused on the question of how well judgement of the attractiveness of each half predicted the perceived attractiveness of the whole face, rather than on the correlation of attractiveness rating for each half with the physical symmetry of the face. The experiment also differed from Scheib et al.'s study in that we asked our participants to rate the attractiveness of female faces rather than asking female participants to rate attractiveness of male faces. Whereas their main interest was the role of symmetry in potential mate choice, our main interest was the perception of facial attractiveness *per se*, without a specific focus on how such perceptual judgements are related to physical health and mate choice.

In Experiment 1, then, participants separately rated attractiveness of the left or right halves of the face and a whole face version. Their ratings for these were compared to each other to see how closely the ratings from the partial and entire face were similar to each other.

We analysed the findings in two ways that address different issues. First, a by participants analysis was used to check that participants did indeed rate the high attractiveness items as more attractive than the low attractiveness items, to see whether part faces were rated as more attractive than whole faces, and to determine whether this applied both to relatively more attractive and to relatively less attractive faces. Second, we used correlational and regression-based analyses of the average ratings for each item to investigate our primary question of the extent to which the attractiveness of the whole face can be predicted from the attractiveness of its parts.

Method

Participants. To establish sufficient power (.80) for detecting the effect of Visible Region in the analysis of rating difference (Experiments 1, 2 and 4A: number of conditions = 3, Experiments 3 and 4B: number of conditions = 5; f = 0.25), the required minimum sample size as estimated by GPower was N = 29 for Experiments 1, 2 and 4A, and N = 25 for Experiments 3 and 4B.

A total of 31 undergraduate students (28 females) aged between 18 and 24 years, *Mdn* = 21, participated in this experiment. All had normal or corrected-to-normal vision. The study was conducted in accordance with the APA's guidelines on the treatment of human participants and was approved by the local ethics committee.

Materials. Face stimuli were developed from a face database from the University of St. Andrews. It contained 702 frontal-view Caucasian faces. Each face came with two versions. One version was the original without cropping, and the other was a cropped version to mask the hair and external features. To eliminate the influence of differences in hairstyle and confine evaluations to facial features *per se*, only the masked version was used in this study. All faces were pre-rated for facial attractiveness, which were based on 19 raters on a 7-point scale. Only female faces were used in this study.

To establish sufficient power (.80) for detecting a medium effect ($R^2 = .25$) of multiple regression with 2 predictors, the required sample size of Experiments 1, 2 and 4A as estimated by GPower was N = 34. Similarly, the required sample size of Experiments 3 and 4B with 4 predictors was N = 42. The required sample size in Experiment 1, 2, 3 and 4 for the correlation with effect size $r = .5$ was N = 29. In addition to determine whether the manipulation in this study affects attractive and unattractive faces equally, we selected the 24 most attractive and 24 least attractive female faces as our stimuli. The mean attractiveness ratings for these two groups of faces (4.11 and 2.23, SD = 0.51, 0.49 respectively) were significantly different from each other ($p < .001$). The use of a total set of 48 items is at the upper limit of many previous studies of perception of facial attractiveness (Rhodes, 2006) and, together with the number of participants tested, led to 1,488 observations per condition, which is consistent with Brysbaert and Stevens' (2018) recommendations for mixed effects designs.

The face width of the original images was normalized to 400 pixels, which subtended 16.6°

of visual angle. Each face was also cut vertically along the middle section of the nose and lip brow to create two half-face versions. Figure 1 shows an example of the three versions (A-C). The manipulation resulted in 48 left halves (where we consider the left half as the half of the photograph falling to the viewer's left) and 48 right halves. Along with the 48 whole faces, there were a total of 144 stimuli in the experiment.

Design and procedure. We employed a within-participant design. Two independent variables were Visible Region (Left, Right, Whole), and Level of Attractiveness (attractive vs. unattractive). The dependent variable was the attractiveness rating on a 7-point scale.

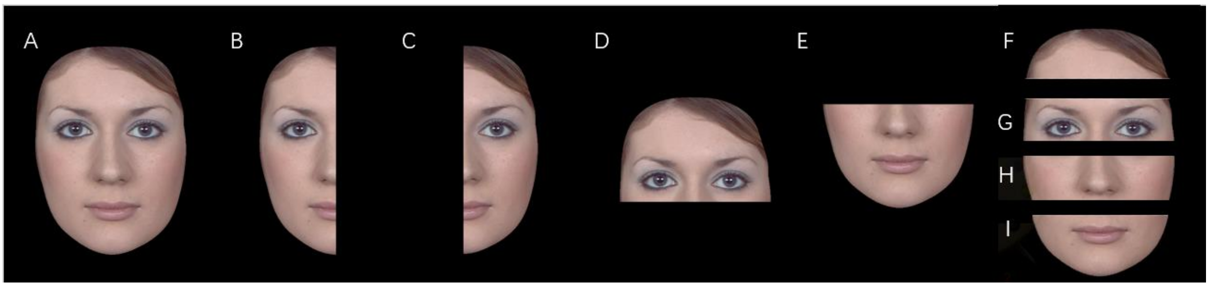

Figure 1. Examples of the stimuli. A: whole face, B: left half, C: right half, D: top half, E: bottom half, F: forehead, G: eyes, H: nose, I: mouth.

Participants were tested individually. Instructions were given on a computer screen. The face stimuli were displayed one at a time in the centre of the screen. A 1-7 scale was shown at the bottom of the screen. Participants were instructed to rate the attractiveness of the face or face part presented using the scale, where 1 represented very unattractive and 7 represented very attractive. If the stimulus was a face part (e.g., left half), the rating task was to judge the attractiveness of face part, but not of any imagined whole face. Participants used a computer mouse to click on the chosen point in the scale. The three conditions of each face identity (left side, right side, whole face) were shown in separate blocks, which were separated by a one-minute break. Each image was randomly assigned to one of the three blocks, such that the three conditions from different persons (e.g., Mary's whole face, Rose's right half and Diana's left half) were intermixed in each block. For example, Mary's whole face, Rose's right half and Diana's left half might appear in the same block, but the three image versions of Mary's face wouldn't appear in the same block. The order of the 48 face identities was randomised for each participant in the first block of trials. Thus, the attractive and unattractive faces were intermixed in each block. The face identities in the second and third blocks followed the same order of presentation as those in the first block, so that the three versions of a face identity were maximally separated from each other in the experiment. This was done to minimise the potential influence of prior memory of the rating given to one version of a face and to encourage a response based on perceived attractiveness of the specific version rather than the memory of a previously given response.

Results

Analysis of Rating Differences. Following the analysis strategy outlined above, we begin with a conventional analysis of the data by participants. Mean ratings across participants as a function of Visible Region (left side, right side, whole face) and Attractiveness (high or low) are shown in Figure 2. A two-way repeated-measures analysis of variance (ANOVA) showed significant main effects of Visible Region, F (2, 60) = 47.12, $p < .001$, $\eta^2_p = .61$, 95% CI of η^2_p [.44, .70], and Attractiveness, F (1, 30) = 311.56, $p < .001$, $\eta^2_p = .91$, 95% CI of η^2_p [.84, .84].

These main effects were qualified by a significant interaction, F (2, 60) = 4.13, p = .02, η^2_p = .12, 95% CI of η^2_p [.002, .26].

To analyze the interaction, we conducted simple effect analyses by Attractiveness. For Attractive faces, there was a significant effect of Visible Region, F (2, 60) = 32.17, p < .001, η^2_p = .52, 95% CI of η^2_p [.32, .63]. Pairwise comparisons with Bonferroni correction showed that the right half was rated more attractive than the left half (p = .02), which was in turn rated more attractive than the whole face (p < .001). For Unattractive faces, the main effect of Visible Region was also significant, F (2, 60) = 31.88, p < .001, η^2_p = .52, 95% CI of η^2_p [.32, .63]. Pair-wise comparisons again showed higher attractiveness ratings for the right half than for the left half (p = .001), which was in turn rated more attractive than the whole face (p < .001).

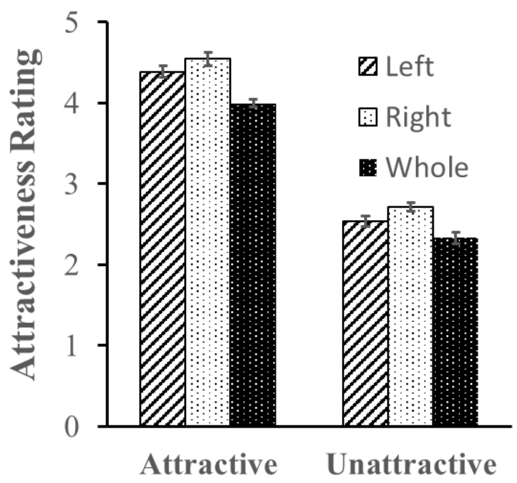

Figure 2. Mean ratings as a function of Visible Region and Attractiveness in Experiment 1. Error bars: one standard error of the means with correction factor appropriate to a within-participant design (O'Brien & Cousineau, 2014).

Both attractive and unattractive faces thus showed a similar pattern of results in the simple effects analyses. As expected, the level of Attractiveness clearly had a substantial overall effect

(Figure 2). As the lower limit of effect size 95% CI (.002) indicates, the interaction between Visible Region and Attractiveness was a relatively small effect which was likely due to a slightly larger difference between the ratings of the right half and the whole face for attractive faces (see Figure 2).

Regression Analysis of Predicting Whole Face Attractiveness from Parts. Analysing differences in mean ratings provides an important context to the findings of the experiment, but our principal interest was in how closely the attractiveness ratings of the face parts were related to the perceived attractiveness of the whole face. To approach this issue, we used regression analysis to assess whether parts could predict whole face.

We calculated the consistency of ratings across different participants; i.e. whether participants agreed with each other sufficiently that their average rating of each item would be meaningful. This was necessary because whilst ratings of the attractiveness of whole faces usually show good inter-rater agreement (Todorov et al., 2015), the same might not hold for face parts.

Usually, studies in this area use Cronbach's alpha as a measure of rater agreement, and this gave high values of .97 for whole faces, .97 for left halves, and .97 for right halves. However, because the appropriateness of using Cronbach's alpha in this way has been questioned (Schmitt, 1996; Kramer et al., 2018), we also calculated absolute agreement based on the intra-class correlation ICC(A, k). Values were again high; .95 for whole faces, .95 for left halves, and .96 for right halves.

We then sought to address our principal aim of determining the extent to which the attractiveness of the whole face can be predicted from the attractiveness of its constituent parts by establishing how strongly the average ratings of a particular face from one condition were correlated with ratings of the same face from a different condition. Correlations between the average rated attractiveness of the part faces and the average rated attractiveness of the whole face images are presented as scatterplots in Figure 3.

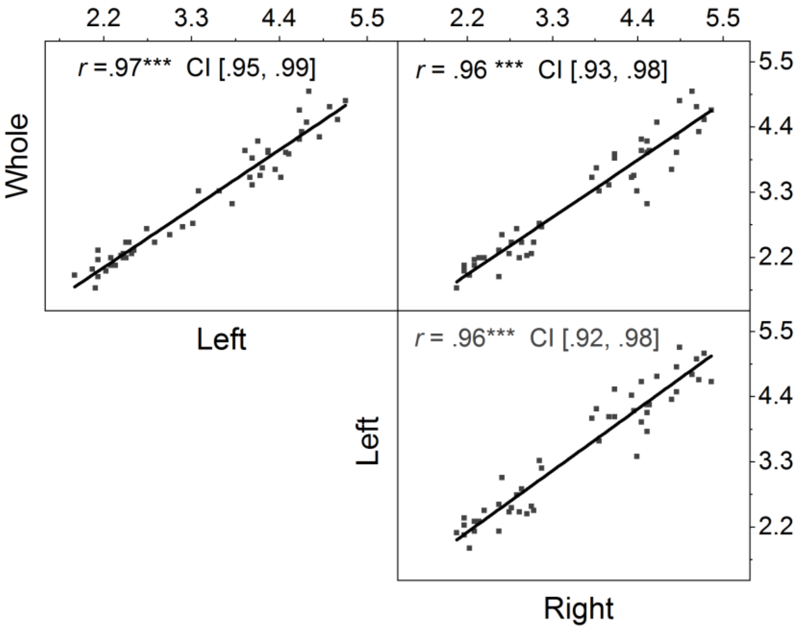

Figure 3. Scatterplots for Correlations Between Average Ratings (across participants) of Corresponding Items from each Condition in Experiment 1. 95% CIs shown in brackets, *** $p < .001$ following Benjamini-Hochberg correction.

These averaged ratings were strongly correlated. The correlation was .96 between ratings of left and right half face, .97 between ratings of left half and whole face, and .96 between ratings of right half and whole face. All were highly significant following application of the Benjamini-Hochberg procedure with a false discovery rate of .05 to correct for multiple comparisons.

We also used a regression analysis to assess how well the attractiveness of the whole face could be predicted from its left or right half. Results with the stepwise method are shown in Table 1, where for convenience they are given alongside the results from later Experiments 2, 3 and 4. The stepwise method applies both forward and backward rules to include or remove variables, and will include only "significant" predictors.

Table 1. Multiple Regression Analyses for Predicting the Whole Face Attractiveness from Face Parts across Experiments 1 to 4.

Experiment	**Visible Region**	**B**	**SE B**	**β**	***t***	***p***
1	Left half	.62	.10	.67	6.20	< .001
	Right half	.29	.10	.32	2.96	.01
2	Top half	.65	.06	.56	10.75	< .001
	Bottom half	.60	.07	.48	9.22	< .001
3	Mouth	.59	.06	.60	10.61	< .001
	Eyes	.38	.05	.40	7.39	< .001
	Forehead	.23	.11	.10	2.18	.04
	Nose			< .01	< 0.01	1.00
4A	Top half	.70	.08	.65	9.02	< .001
	Bottom half	.33	.07	.35	4.87	< .001
4B	Eyes	.59	.08	.62	7.19	< .001
	Mouth	.35	.08	.38	4.45	< .001
	Forehead			-.15	-1.43	.162
	Nose			.28	1.74	.090

For Experiment 1, either the left or the right half of the face alone predicted the whole face attractiveness, F's (1, 46) = 814.41 and 507.97, p's < .001, adjusted R^2's were .95 and .92, 95% CI of R^2 were [.90, .97], [.85, .95], respectively. When the two halves were combined, prediction for the whole face attractiveness improved, and together they explained a greater proportion of variance in whole face attractiveness, adjusted $R^2 = .95$, $F(2, 45) = 479.99$, $p < .001$, 95% CI of R^2 [.91, .97].

Discussion

Results in this experiment showed that either the left or right half of a face alone can predict the attractiveness of the whole face. Moreover, the rated attractiveness of one half of the face was highly correlated with the attractiveness of the other half ($r = .96$). Since faces are only partly asymmetric (Acheib et al., 1999; Rhodes, 2006), this latter result is perhaps unsurprising, but it offers a useful context to findings that will be reported in Experiments 1 and 2, where the parts tested are no longer even approximate mirror images of each other.

Both attractive and unattractive faces were rated more attractive when only half of the face was shown, relatively to the whole face. The effect of this was slightly bigger for attractive faces. For both attractive and unattractive faces the right side of faces was rated as more attractive.

The fact that the left or right sides presented in isolation were judged as being slightly more attractive than the whole faces is consistent with ideas about the importance of bilateral symmetry (Scheib et al., 1999), since the whole faces will be slightly asymmetric but when presented with only a left or right side it will be reasonable to overestimate symmetry by assuming that the other side would be much the same. However, this strategy should not work when the top or bottom of a face is used to estimate the features of the other half because there are not symmetrical features across the vertical dimension of the face. Can attractiveness of the top half or the bottom half predict attractiveness of the whole face or of the other part face without using any symmetrical information across the two halves? This question was examined in the next experiment.

Experiment 2

The purpose of this experiment was to examine whether the main findings of Experiment 1 could be replicated when the visible half did not contain any symmetrical information of the invisible half. Participants rated attractiveness of the top and bottom halves of a face and the results were compared to each other and to the rated attractiveness of the whole face.

Method

Participants. A total of 31 university students (*Mdn* age 20 years, range 18-27 years, 30 females) participated in this experiment. All had normal or corrected-to-normal vision.

Materials. The same 24 attractive and 24 unattractive face identities as Experiment 1 were used. Again, each original face was used to create two other versions, but this time either the bottom or top the half of the face was masked. Figure 1 (examples D and E) illustrates this manipulation.

Design and procedure. These were also identical to Experiment 1, except that the Left and Right Half of independent variable Visible Region were replaced with Top and Bottom Half.

Results

Analysis of Rating Differences.

The mean attractiveness rating results are show in Figure 4. ANOVA showed significant main effects of Visible Region, $F(2, 60) = 4.47$, $p = .02$, $\eta^2_p = .13$, 95% CI of η^2_p [.01, .27], and Attractiveness, $F(1, 30) = 282.19$, $p < .001$, $\eta^2_p = .90$, 95% CI of η^2_p [.82, .93]. These effects were qualified by a two-way interaction, $F(2, 60) = 15.09$, $p < .001$, $\eta^2_p = .34$, 95% CI of η^2_p [.14, .48].

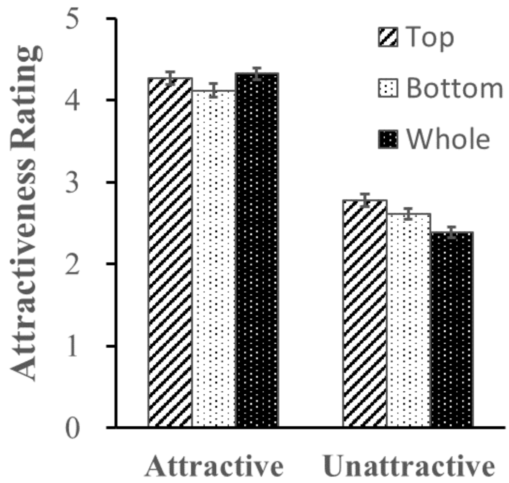

Figure 4. Mean ratings as a function of Visible Region and Attractiveness in Experiment 2. Error bars: one standard error of the means with correction factor appropriate to a within-participant design (O'Brien & Cousineau, 2014).

To identify the source of the interaction, we conducted simple effects analyses separately for the two Attractiveness conditions. For Attractive faces, there was a significant main effect of Visible Region, $F(2, 60) = 3.19$, $p = .05$, $\eta^2_p = .10$, 95% CI of η^2_p [0, .23]. Pairwise comparisons showed no reliable differences among the conditions. For Unattractive faces, there was also a main effect of Visible Region, $F(2, 60) = 14.39$, $p < .001$, $\eta^2_p = .32$, 95% CI of η^2_p

[.13, .47]. Pairwise comparisons showed that both Top and Bottom were rated more attractive than the whole face. No difference was found between ratings of Top half and Bottom half. However, the difference among Top half, Bottom half and whole face may be small for attractive faces as the small lower limit of effect size 95% CI (0) indicates.

Regression Analysis of Predicting Whole Face Attractiveness from Parts. We again calculated inter-rater agreement using Cronbach's alpha and the intra-class correlation ICC(A, k). Values were again high; Cronbach's alpha .98 for whole faces, .97 for top halves, and .96 for bottom halves, ICC(A, k) .97 for whole faces, .95 for top halves, and .95 for bottom halves.

Turning to the key issue of how well the attractiveness of the face could be predicted from its constituent parts, scatterplots for the correlations between the average rated attractiveness of the part faces and the average rated attractiveness of the whole face images are presented in Figure 5A.

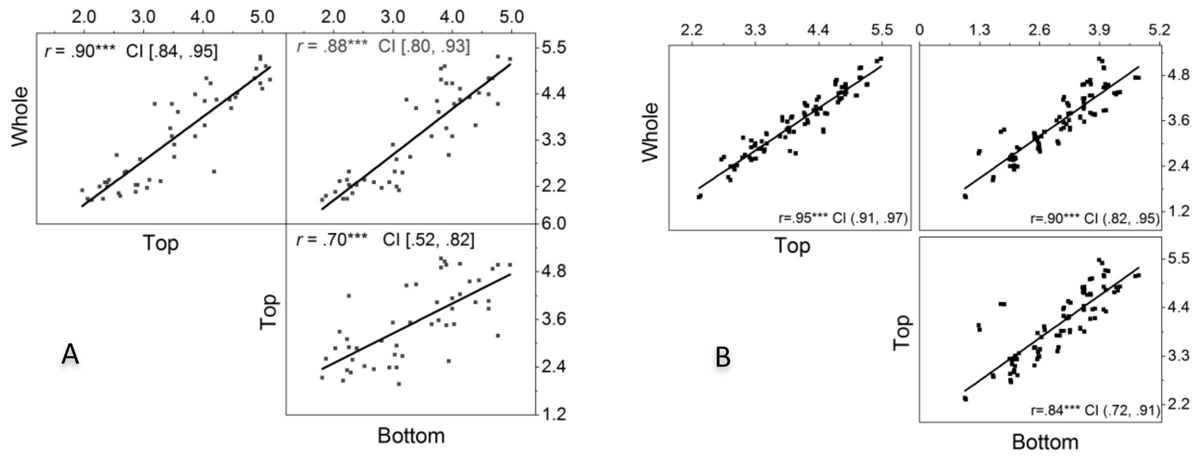

Figure 5. Scatterplots for Correlations Between Average Ratings (across participants) of Corresponding Items from each Condition in Experiment 2 (panel A) and Experiment 4A (panel B). 95% CIs shown in brackets, *** $p < .001$ following Benjamini-Hochberg correction.

Figure 5A shows that the average rating of top-half faces was strongly correlated with the rating of the bottom-half, r = .70, although the strength was lower than the near perfect left-right half correlation found in Experiment 1 (r = .96). Ratings for both top and bottom halves were highly correlated with ratings for the whole face, r's = .91, and .88, respectively, although the strengths were again lower than for left/right-whole correlations, which were .97 and .96, respectively in Experiment 1. All of these correlations were highly significant following application of the Benjamini-Hochberg procedure with a false discovery rate of .05 to correct for multiple comparisons.

A regression analysis showed that either the top or the bottom half of the face alone predicted the whole face attractiveness, F's (1, 46) = 207.03 and 159.01, p's < .001, adjusted R^2 were .81 and .77, 95% CI of R^2 were [.69, .89], [.62, .86], respectively. When combined, both the top half and the bottom half significantly predicted the whole face attractiveness (see Table 1), and they together explained a significant proportion of variance in whole face attractiveness, adjusted R^2 = .93, $F(2, 45) = 335.11$, $p < .001$, 95% CI of R^2 [.88, .96].

Discussion

Results in this experiment showed that the top and bottom half of the face alone provide reasonably accurate prediction of the whole face attractiveness. Although the predictions made from these single parts were less accurate than from the left/right half of the face, the prediction from the top and bottom halves increased significantly when both parts were taken into consideration.

Whilst the correlation between rated attractiveness of the face parts themselves (r = .70) was lower than that between the top half and the whole face (r = .91) or the bottom half and the whole face (r = .88), it none the less accounted for 49% of the variance in ratings even though the top and bottom halves are not mirror images of each other.

Unlike Experiment 1, where left/right parts were rated more attractive than the whole face, top/bottom parts in this experiment were not clearly different from the whole face for the attractive faces, although both parts were rated more attractive than the whole face for unattractive faces. Despite some of these differences, both Experiments 1 and 2 showed how well attractiveness of the whole face can be estimated from just half of the image, whether it was from left/right half, or from top/bottom half. In Experiment 3 we further investigated whether prediction about the whole face attractiveness could be made from even smaller regions of the face.

Experiment 3

In this experiment, we used a modified version of Santos and Young’s (2011) 'isolated region' method to segment faces horizontally into four local regions: forehead, eyes, nose, and mouth. Each of the four regions was exposed through a rectangular opening while rest of the face was masked (see Figure 1, examples F-I). Given the importance of the eye and mouth regions in perception of attractiveness, we predicted that these regions would be more important predictors for whole face attractiveness relative to other regions.

Method

Participants. Fifty-six university students (*Mdn* age 21 years, range 19-29 years, 40 females) participated in the experiment. All had normal or corrected-to-normal vision.

Materials. The same faces as Experiments 1 and 2 were used in this experiment. However the face parts were segmented differently. Each of the 48 faces was segmented into four regions: the forehead, eyes, nose, and mouth region. Along with the whole face, this resulted in a total of 240 images (48 × 5). Example stimuli of the four segmented face regions are given in Figure 1 (examples F-I).

Design and procedure. This was a within-participant design. The two factors were Visible Region (forehead, eyes, nose, mouth, and the whole face) and Attractiveness (attractive vs. unattractive).

The procedure was essentially the same as in Experiments 1 and 2, except that two additional blocks of 48 trials were added to accommodate the five conditions of the visible region variable. As in the first two experiments, each trial within one block showed a unique face image, and a different image version of each face was presented in a different block. The randomisation procedure was also the same as before, hence all conditions were mixed in each block. The 48 face identities with five visible region conditions led to a total of 240 trials. All other aspects of the procedure and task were identical to the first two experiments.

Results

Analysis of Rating Differences. The mean attractiveness ratings of this experiment are shown in Figure 6. The main effects were significant for both Visible Region, $F(4, 224) = 46.05$, $p < .001$, $\eta^2_p = .45$, 95% CI of η^2_p [.35, .52] and Attractiveness, $F(1, 56) = 441.52$, $p < .001$, $\eta^2_p = .89$, 95% CI of η^2_p [.83, .92]. There was also a significant a two-way interaction, $F(4, 224) =$

97.18, $p < .001$, $\eta^2_p = .63$, 95% CI of η^2_p [.56, .68].

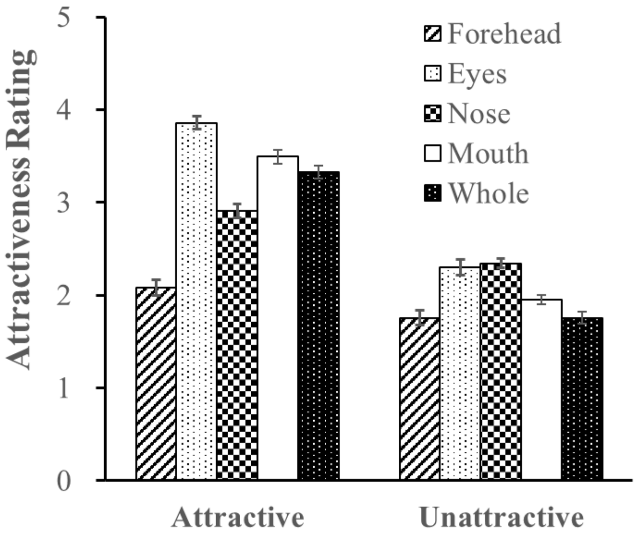

Figure 6. Mean ratings as a function of Visible Region and Attractiveness in Experiment 3. Error bars: one standard error of the means with correction factor appropriate to a within-participant design (O'Brien & Cousineau, 2014).

Simple effects analyses showed significant effect of Visible Region for both attractive faces, $F(4, 224) = 82.36$, $p < .001$, $\eta^2_p = .60$, 95% CI of η^2_p [.51, .65], and unattractive faces, $F(4, 224) = 20.30$, $p < .001$, $\eta^2_p = .27$, 95% CI of η^2_p [.16, .35]. Pairwise comparisons with Bonferroni correction were used to further analyze these main effects. For attractive faces, the eyes were rated significantly more attractive than everything else (p's $\leq .004$), followed by the mouth and the whole face, which were not rated differently from each other ($p = .92$), but were both rated more attractive than the nose and the forehead (p's $< .001$). The forehead was rated significantly lower than all other regions (p's $< .001$). For unattractive faces, pairwise comparisons yielded fewer significant differences. Ratings were comparable for the eyes and the nose (p = 1.00), but both were rated more attractive than for forehead, the mouth, and the whole face (p's ≤ .006). While the mouth was rated more attractive than the whole face (p = .002), it was not rated differently from the forehead (p = .13), which was also rated the same as the whole face (p = 1.00).

The two-way interaction can be also seen to involve varying degrees of difference between ratings of the corresponding face region in the two attractiveness conditions. For example the ratings for the forehead region of attractive and unattractive faces showed a much smaller difference (0.33) than other regions, such as between ratings for the eyes (1.56) or mouth region (1.54). Interestingly, however, paired t tests for these comparisons revealed that even the smallest difference (i.e., the forehead region) showed a statistically higher rating for the part from attractive faces, $t(56) = 6.84$, $p < .001$, Cohen's D = 0.37, 95% CI of D [0.25, 0.51].

Regression Analysis of Predicting Whole Face Attractiveness from Parts. We again calculated inter-rater agreement using Cronbach's alpha and the intra-class correlation ICC(A, k). Values were again high; Cronbach's alpha .98 for whole faces, .93 for the forehead region, .98 for eye region, .92 for nose region, and .98 for mouth region, ICC(A, k) .97 for whole faces, .87 for the forehead region, .96 for eyes, .84 for the nose region, and .97 for the mouth.

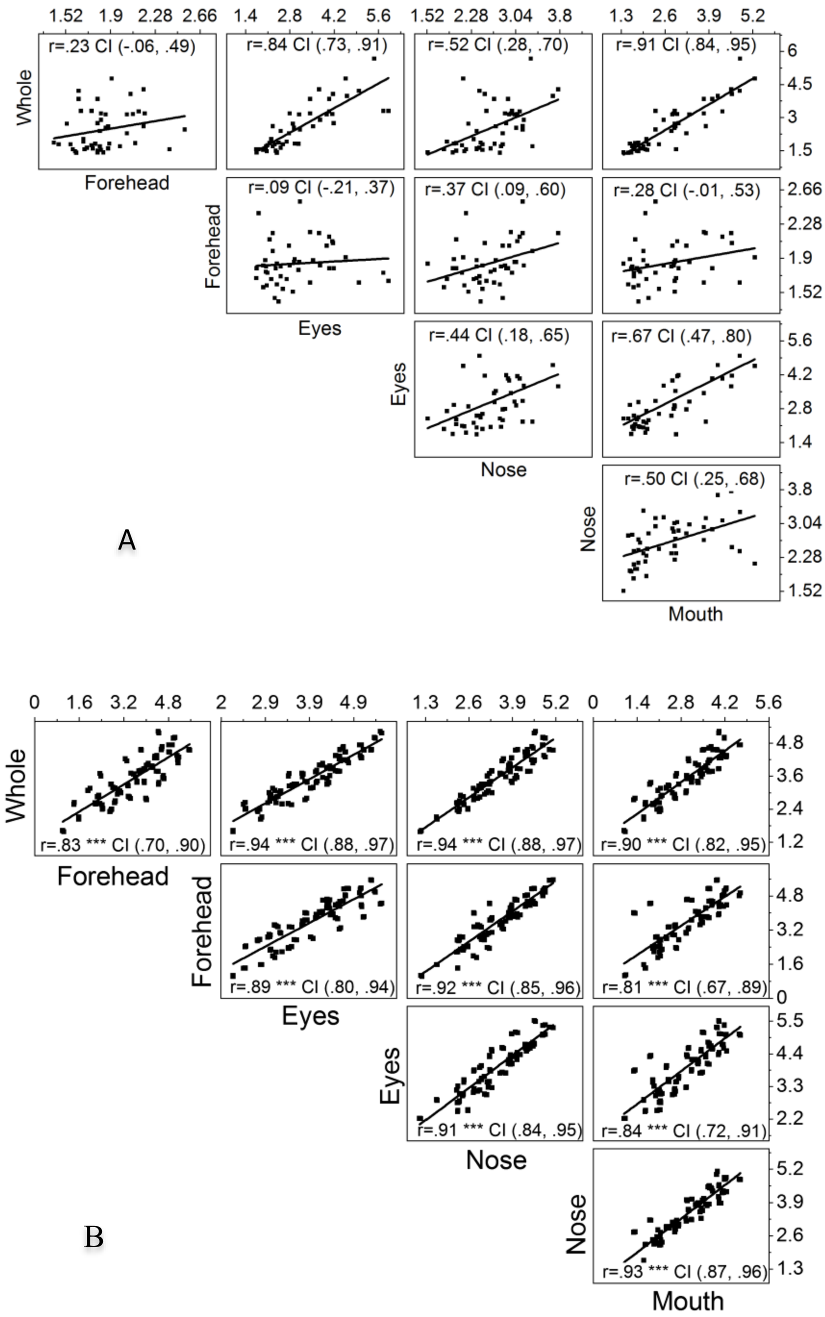

Figure 7. Scatterplots for Correlations Between Average Ratings of Whole Face, Forehead, Eyes, Nose and Mouth Regions in Experiment 3 (panel A) and Experiment 4B (panel B). 95% CIs shown in brackets, ***p* < .01, *** *p* < .001 following Benjamini-Hochberg correction.

Figure 7A shows that the average ratings of face parts and whole face were often strongly correlated, where significance levels are based on application of the Benjamini-Hochberg procedure with a false discovery rate of 0.05 to correct for multiple comparisons. The scales for rated attractiveness of the forehead and nose regions have been adjusted in Figure 7A to reflect the limited (and relatively low) range of rated attractiveness of these regions.

As Figure 7A shows, the average rating score for the whole face was positively correlated with average ratings of mouth, eyes, nose (*r*'s = .91, .84, .52, respectively, *p*'s < .001), but not with forehead ($r = .23$, $p = .124$). In addition, the ratings of the attractiveness of all of the face regions except the forehead were significantly correlated with each other.

A regression analysis showed that the mouth, eyes, or nose alone predicted the whole face attractiveness, *F*'s (1, 46) = 218.65, 107.47, 17.35, *p*'s < .001, adjusted R^2's were .82, .69, .26, 95% CI of R^2 were [.70, .89], [.51, .81], [.07, .47], respectively. However the forehead didn't significantly predict the whole face attractiveness, $F(1, 44) = 2.46$, $p = .124$, adjusted R^2's were .05, 95% CI of R^2 were [0, .23] Based on the 95% CI of effect size, the nose or forehead might contribute to the attractiveness of whole face, but the effect might be smaller. When combined, only the mouth, the eyes and the forehead significantly predicted the whole face attractiveness (see Table 1), and they together explained a significant proportion of variance in whole face attractiveness, adjusted $R^2 = .92$, $F(3, 44) = 191.75$, $p < .001$, 95% CI of R^2 [.86, .95].

Discussion

The results of this experiment show that the isolated eyes, mouth, nose and forehead regions can each predict the attractiveness of whole faces, but some of these face regions provide better estimates of whole face attractiveness than others, as indicated by R^2. The mouth and the eye regions were the most effective predictors ($R^2 = .82, .69$), whereas in contrast the forehead provided less information about the whole face attractiveness ($R^2 = .24$). However, it is worth noting that even the rating scores for the forehead of attractive faces were significantly higher than those of unattractive faces, and that every region of the face was correlated to some degree both with overall attractiveness of the whole face and with the attractiveness of each of the other regions.

Experiment 4

In the previous three experiments, we used only female faces, some of which might use makeup which might have influenced our findings since cosmetics can be used to emphasise sex differences in facial contrast (Russell, 2009). To tackle this potential problem, male faces from a publicly available face database were used in Experiment 4, since males are less likely to use to makeup. Experiments 1-3 also used stimuli selected only from the attractive or unattractive ends of the continuum of facial attractiveness. To explore the possibility that this bimodal distribution might also have influenced the findings of Experiments 1-3, Experiment 4 used 90 male faces with attractiveness scores falling along a normal continuum. Considering the relatively low variance of left, right and whole faces, only the conditions of Experiment 2 and 3 were replicated in Experiment 4. Experiment 2 was replicated in Experiment 4A, whereas Experiment 3 was replicated in Experiment 4B.

Method

Participants. Forty-one Chinese university students (*Mdn* age 20.5 years, SD=2.2, range 18-25 years, 24 females) participated in the experiment. All had normal or corrected-to-normal vision.

Materials. We adopted stimuli from the CUHK Face Sketch Database (CUFS, Wang & Tang, 2009, http://mmlab.ie.cuhk.edu.hk/archive/facesketch.html). This database has 134 male faces of students of the Chinese University of Hong Kong (CUHK), and 90 faces were selected after excluding those older students and non-Chinese ethnicity. Each of the 90 faces were segmented into bottom and top halves (Experiment 4A) or forehead, eyes, nose, and mouth regions (Experiment 4B). Along with the whole face, this resulted in a total of 630 images (90 × 7).

Design and procedure. To replicate Experiments 2 and 3, there were two within-participant designs with one repeated-measure independent variable: Experiment 4A, Visible Region (bottom half, top half, and whole face); Experiment 4B, Visible Region (forehead, eyes, nose, mouth, and whole face).

The procedure was essentially the same as in Experiments 2 and 3, except that the stimuli were faces from the CUFS database.

Results

The attractiveness ratings of this experiment are shown in Table 1. The attractiveness scores of the whole faces ranged from 1.85 to 4.76 with a mean of 3.53 (SD = 0.61).

Table 2. Ratings as a function of Visible Region in Experiment 4 (N_{item}= 90).

Visible region	Minimum	Maximum	Mean	Std. Error
Whole	1.85	4.76	3.53	.06
Bottom	1.90	4.39	3.00	.04
Top	2.32	5.98	4.00	.08
Forehead	2.27	4.51	3.54	.04
Eyes	1.88	5.56	3.96	.08
Nose	2.71	4.39	3.42	.04
Mouth	2.10	4.39	2.98	.04

We again calculated inter-rater agreement using Cronbach's alpha and the intra-class correlation ICC(A, k). Values were again high. In Experiment 4A, Cronbach's alpha .99 for whole faces, .99 for bottom half, .98 for top half, ICC(A, k) .98 for whole faces, .99 for bottom, .97 for top. In Experiment 4B, Cronbach's alpha .99 for the forehead region, .99 for eye region, .99 for nose region, and .99 for mouth region, ICC(A, k) .98 for whole faces, .99 for the forehead region, .98 for eyes, .99 for the nose region, and .99 for the mouth.

Analysis of Rating Differences -- Experiment **4A:**

The main effects were significant for Visible Region, $F(2, 178) = 97.45$, $p < .001$, $\eta^2_p = .52$, 95% CI of η^2_p [.42, .60]. Pairwise comparisons showed that top halves were more attractive than the whole face, and the whole face was more attractive than bottom halves (p’s < .001).

Analysis of Rating Differences -- Experiment **4B:**

The main effects were significant for Visible Region, $F(4, 356) = 56.65$, $p < .001$, $\eta^2_p = .39$, 95% CI of η^2_p [.32, .44]. Pairwise comparisons showed that eyes were more attractive than other parts and whole face (p’s < .001); the whole face and forehead were similar in attractiveness (p = .885), but both were more attractive than nose and mouth (p’s <= .05); and nose more attractive than mouth ($p < .001$).

Regression Analysis of Predicting Whole Face Attractiveness from Parts -- Experiment 4A

Figure 5B shows that the average rating of top-half faces was strongly correlated with the rating of the bottom-half, $r = .84$, and ratings for both top and bottom halves were highly correlated with ratings for the whole face, r’s = .95, and .90, respectively.

A regression analysis showed that either the top or the bottom half of the face alone predicted the whole face attractiveness, F’s (1, 88) = 365.00 and 159.91, p’s < .001, adjusted

R^2 were .90 and .81, 95% CI of R^2 were [.85, .93], [.72, .87], respectively. When combined, both the top half and the bottom half significantly predicted the whole face attractiveness (see Table 1), and they together explained a significant proportion of variance in whole face attractiveness, adjusted $R^2 = .94$, $F(2, 87) = 300.64$, $p < .001$, 95% CI of R^2 [.90, .96].

Regression Analysis of Predicting Whole Face Attractiveness from Parts - Experiment 4B

As Figure 7B shows, the average rating scores for the whole face were positively correlated with average ratings of mouth, eyes, nose and forehead. In addition, the ratings of the attractiveness of all of the face regions were significantly correlated with each other.

A regression analysis showed that the mouth, eyes, nose or forehead alone predicted the whole face attractiveness, F's (1, 88) = 167.01, 280.65, 279.22, and 83.91, p's < .001, adjusted R^2's were .81, .88, .87 and .68, 95% CI of R^2 were [.72, .87], [.81, .92], [.81, .91], [.55, .77], respectively. When combined, only the mouth and eyes significantly predicted the whole face attractiveness (see Table 1), and they together explained a significant proportion of variance in whole face attractiveness, adjusted $R^2 = .92$, $F(2, 87) = 217.98$, $p < .001$, 95% CI of R^2 [.87, .94].

Discussion

In line with Experiments 2 and 3, the results of this experiment showed that the isolated eyes, mouth, nose and forehead regions as well as top and bottom half can each predict the attractiveness of whole faces. Again, the top and bottom half can predict the whole face well. Of the parts, the mouth and the eye regions were still the most effective predictors. The nose was also an effective predictor, and the forehead also provided some information about the whole face attractiveness, but their contributions were largely overlapping with the eyes and mouth. The rating scores for the forehead were also higher than Experiment 3.

In sum, Experiment 4 replicated the patterns of results of Experiments 2 and 3 with a new set of faces without makeup and using the full range of facial attractiveness.

General Discussion

This study tested whether a part of a face can contain enough information to predict the attractiveness of the whole face. The results support this hypothesis: attractiveness ratings of the left and right halves of the face (Experiment 1) or the top and bottom halves of a face (Experiments 2 and 4A) were highly correlated with the ratings for the whole face, and even ratings for much smaller regions of the face (eyes, nose, mouth, forehead) were correlated with the ratings of the whole face (Experiments 3 and 4B). Overall, these findings demonstrate substantial natural covariation between different parts of a face, making it possible to infer attractiveness from a relatively small part of the face.

Nonetheless, our results also show that not all parts of the face were equally predictive of the whole face attractiveness. Predictions for the whole face were the best from the left or right half of the face (Experiment 1) followed by the top or bottom half of the face (Experiments 2 and 4A), which was better than the mouth and the eyes (Experiments 3 and 4B). The eye and mouth regions were themselves more effective relative to the nose and the forehead regions. However, it is worth noting that the use of a smaller face region never resulted in that region being uninformative.

The inference of covariation in the cues that convey attractiveness is substantially strengthened by the observation that not only did the ratings of part faces correlate with rated attractiveness of the whole face, but ratings of the different parts invariably correlated significantly with each other. However, we do not seek to deny that multiple cues are used simultaneously to determine attractiveness, as is demonstrated through findings of holistic perception (e.g. Todorov et al., 2010; Lin & Zhou, 2021). Our point is only that there is substantial natural covariation between these cues.

How might we account for this covariation in the attractiveness of different face regions? Why should someone who has attractive eyes be likely also to have an attractive mouth? Dominant accounts of facial attractiveness suggest a link to sexual selection via perceived health and 'mate quality'. From this perspective, substantial covariation would be expected between different cues that can underlie perceived attractiveness (Fink & Penton-Voak, 2002). Whilst this is consistent with our findings, and it is impressive that it predicts what we observed, it is by no means the only possibility. Attractiveness judgements might be in part aesthetic without being mixed with sexual interests, and we note that variables such as averageness affect the perceived aesthetic attractiveness of many stimuli other than faces (Halberstadt & Rhodes, 2000). We note too that observers of either sex are able to judge the attractiveness of same-sex faces; an observation which requires the additional assumption that we need to judge the attractiveness of potential rivals to fit the evolutionary perspective. Moreover, our findings are consistent with other theoretical positions, such as the hypothesis that attractiveness is correlated with averageness (Langlois et al., 1994), since a truly 'average' face will presumably have an average nose, average eyes, an average mouth and so on.

Whilst the best way to explain the covariation in attractiveness we observed across different face regions remains open, then, it is clearly an interesting phenomenon. However, because we used face photographs as stimuli, it is important to note that different images of the same face can be highly variable in perceived attractiveness due to differences in lighting, pose, expression and so on; in fact the differences between images of the same face can be as large as the differences in overall attractiveness between the faces of different individuals (Jenkins et al., 2011; Sutherland et al., 2017; Todorov & Porter, 2014). Here, we minimised the impact of image variability by using an existing set of face photographs taken under standardised conditions.

Also worth noting is that our correlation and regression analyses are based on ratings of images that were averaged across participants. Whilst it is known that different individuals show substantial agreement in their ratings of attractiveness, reflecting an underlying 'shared taste', there are also individual differences that can reflect more personal 'private taste' as to what is attractive (Germine et al., 2015; Hehman et al., 2017). Averaging ratings in this way eliminates much of the individual differences between raters that constitute their private taste, so that the findings then largely reflect the shared taste as to what is attractive that is common to most raters.

Our findings call for reinterpretations of some main theories of facial attractiveness. For example, symmetry and averageness have been considered as main criteria for facial attractiveness. Yet in our stimuli that showed only a part of a face, information about symmetry and averageness of the whole face was largely absent. This, however, did not seriously hamper judgments of the attractiveness. Our results show that other factors than symmetry alone play a role in judgment of face parts. Even though the left/right half of a face contains little information of symmetry, participants had no problem in judging its attractiveness and their estimates of the attractiveness of half faces were strongly related to the attractiveness of the whole face images. Interestingly, ratings for both left and right halves were actually higher than for the whole face (see Figure 2). This small yet significant dip of perceived attractiveness could represent the role of bilateral symmetry in attractiveness judgment. One possibility is that the inference of the whole face attractiveness from the left/right half is based on the assumption that the features on the visible half mirror the invisible half. When the whole face is shown, the left-right asymmetries that are not present in the left or right half-face image would overwrite the assumption, which result in a slightly compromised attractiveness judgment.

The findings that rating scores for part faces were highly correlated with ratings of the whole faces and with each other implies two things. First, inferences about the attractiveness of the face from a face part can rely on a lot of other types of information than bilateral symmetry. Second, there is substantial covariation across different cues to attractiveness.

It is known that factors such as the skin colour and texture are important factors in assessing facial attractiveness (Little et al., 2011), possibly because these serve as cues to such other important characteristics as physical health (Rhodes, 2006). Skin texture and colour are likely to be correlated across different parts of a face. The brain may then use this knowledge when a part of the face is being used to infer the whole face attractiveness. This correlated nature may generalise beyond more distributed features such as skin tone and texture. In Experiment 3, the eye and mouth regions clearly predicted most variance, but it is interesting that the relatively 'featureless' forehead region was of any use at all. Perhaps it is informative to the extent that skin tone offers an index of potential attractiveness.

A localised facial feature such as the mouth may also correlate with the level of attractiveness of the other features such as the eyes and skin texture on the same face. The participants appear to act as if they make the assumption that if one region of a face is of a certain degree of attractiveness, then the other (invisible) regions should have a similar level. This assumption would allow for inferences about the occluded regions of the face. The high correlations between parts of a face in our data demonstrate that such an assumption has solid ecological validity. In fact, our finding that face parts can predict each other may be more interesting than the finding of a part predicting the whole, because it shows that a part can predict another face region even when it is not included in that region. Perhaps it is this kind of covariation that the visual system utilises when a visible region of the face is limited by occlusion. Covariation has been used in evolutionary accounts of facial attractiveness, and it is worth noting that perceptual mechanisms could employ covariation in nature to infer an unseen region of the face from the visible region. Moreover, useful covariation between cues is also evident in other areas of face perception (Young, 2018).

Rating attractiveness based on face parts in these experiments tended to show an overestimation bias relative to the whole face (Experiments 1, 3 and 4). Only rating the top or the bottom half for the attractive faces in Experiment 2 showed exactly the same as the whole face. This means that participants tended to see parts as being slightly more attractive in the absence of information from the other regions, as has also been noted in previous studies (Liu & Chen, 2018; Saegusa & Watanabe, 2016; Orghian & Hidalgo, 2020). This consistent finding has been interpreted as indicating a positivity bias in which missing information may be assumed to come from an attractive face (Orghian & Hidalgo, 2020). This bias will be corrected as greater face regions are exposed, so that any evidence of imperfection can accumulate. Consequently, the new information may correct the assessment based on a fragment. This is also consistent with the time course of attractiveness evaluation, in which a brief exposure to a face can create higher attractiveness rating than a longer exposure, again presumably due to more complete information with longer exposure duration revealing more imperfections (Liu & Chen, 2018; Saegusa & Watanabe, 2016).

Our results show that local properties in different regions of the face can form the basis for reasonable inferences about facial attractiveness of partially occluded faces. They complement findings that multiple cues combine to determine the overall perceived attractiveness of a facial image (Holzleitner et al., 2019; Said & Todorov, 2011; Sutherland et al., 2013; Vernon et al., 2014) with a very different procedure and suggest that the visual system is able to exploit natural cue covariation to support the rapid appraisal of physical attractiveness.

References

Abbas, Z. A., & Duchaine, B. (2008). The role of holistic processing in judgments of facial attractiveness. *Perception, 37*, 1187-1196.

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion, 6*, 269-278.

Baudouin, J. Y., and Tiberghien, G. (2004). Symmetry, averageness and feature size in the facial attractiveness of women. *Acta Psychologica, 117*, 313-332. doi: 10.1016/j.actpsy.2004.07.002

Bronstad, P. M., Langlois, J. H., Russell, R. (2008). Computational models of facial attractiveness judgments. *Perception, 37*, 126-142.

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1(1),* 9, 1-20, DOI: https://doi.org/10.5334/joc.10

Cunningham, M. R. (1986) Measuring the physical in physical attractiveness: Quasi-experiments on the sociobiology of female facial beauty. *Journal of Personality and Social Psychology, 50*, 925-935.

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*, 285-290.

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979) Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception, 8*, 431-439.

Fink, B., & Penton-Voak, I. (2002). Evolutionary psychology of facial attractiveness. *Current Directions in Psychological Science, 11*, 154-158.

Fink, B., Grammer, K., & Matts, P. J. (2006). Visible skin color distribution plays a role in the perception of age, attractiveness, and health in female faces. *Evolution and Human Behavior*, 27, 433-442.

Gill, D. (2017). Women and men integrate facial information differently in appraising the beauty of a face. *Evolution and Human Behavior*, 38, 756-760.

Gao, Z., & Bentin, S. (2011). Coarse-to-fine encoding of spatial frequency information into visual short-term memory for faces but impartial decay. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1051-1064.

Germine, L., Russell, R., Bronstad, P. M., Blokland, G. A. M., Smoller, J. W., Kwok, H., … Wilmer, J. B. (2015). Individual aesthetic preferences for faces are shaped mostly by environments, not genes. *Current Biology, 25*(20), 2684-2689.

Halberstadt, J., & Rhodes, G. (2000). The attractiveness of nonface averages: implications for an evolutionary explanation of the attractiveness of average faces. *Psychological Science, 11*, 285-289.

Hehman, E. A., Sutherland, C. A. M., Flake, J. K., & Slepian, M. L. (2017). The unique contributions of perceiver and target characteristics in person perception. *Journal of Personality and Social Psychology, 113(*4), 513-529.

Holzleitner, I. J., Lee, A. J., Hahn, A. C., Kandrik, M., Bovet, J., Renoult, J. P., Simmons, D., Garrod, O., DeBruine, L. M., & Jones, B. C. (2019). Comparing theory-driven and data-driven attractiveness models using images of real women's faces. *Journal of Experimental Psychology: Human Perception and Performance*, *45*, 1589-1595.

Huang, Y., & Rao, R. P. N. (2011). Predictive coding. *WIREs Cognitive Science, 2*, 580-593.

Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition, 121*, 313-323.

Johnston, P., Overell, A., Kaufman, J., Robinson, J., & Young, A. W. (2016). Expectations about person identity modulate the face-sensitive N170. *Cortex, 85*, 54-64.

Johnston, P., Robinson, J., Kokkinakis, A., Ridgeway, S., Simpson, M., Johnson, S., Kaufman, J., & Young, A. W. (2017). Temporal and spatial localization of prediction-error signals in the visual brain. *Biological Psychology, 125*, 45-57.

Johnston, V.S., & Franklin, M. (1993). Is beauty in the eye of the beholder? *Ethology and Sociobiology, 14* (3), 183-199.

Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2004). When facial attractiveness is only skin deep. *Perception, 33*(5), 569-576.

Kramer, R. S. S., Mileva, M., & Ritchie, K. L. (2018). Inter-rater agreement in trait judgements from faces. *PLoS ONE, 13(8)*, e0202655.

Langlois, J. H., Roggman, L. A., & Musselman, L. (1994). What is average and what is not average about attractive faces? *Psychological Science, 5*, 214-220.

Lin, J., & Zhou, G. (2021). Chinese Aesthetic Mask: Three Forehead and Five Eyes—Holistic Processing and Facial Attractiveness. *Perception,* 50(6), 540–554.

Liu, C. H., & Chen, W. (2018). The boundary of holistic processing in the appraisal of facial attractiveness. *Royal Society Open Science, 5*: 171616. DOI: 10.1098/rsos.171616.

Little A. C., Jones, B. C., & DeBruine, L. M. (2011). Facial Attractiveness: Evolutionary Based

Research. *Philosophical Transactions of the Royal Society B, 366*, 1638-1659.

Locher, P., Unger, R., Sociedade, P., & Wahl, J. (1993). At first glance: Accessibility of the physical attractiveness stereotype. *Sex Roles, 28*, 729-743.

O'Brien, F., & Cousineau, D. (2014). Representing Error bars in within-subject designs in typical software packages. The Quantitative Methods for Psychology, 10 (1), 56-67.

Orghian, D., & Hilgado, C. A. (2020). Humans judge faces in incomplete photographs as physically more attractive. *Scientific Reports 10*, 110.

Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., Henzi, S. P., Castles, D. L., & Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature, 394,* 884-887.

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57*, 199-226.

Rubenstein, A. J., Langlois, J. H., & Roggman, L. A. (2002). What makes a face attractive and why: The role of averageness in defining facial beauty. In G. Rhodes & L. A. Zebrowitz (Eds.), *Advances in Visual Cognition, Vol. 1. Facial Attractiveness: Evolutionary, Cognitive, and Social Perspectives* (pp. 1-33). Westport, CT, US: Ablex Publishing.

Russell, R. (2009). A sex difference in facial contrast and its exaggeration by cosmetics. *Perception, 38*, 1211-1219.

Russell, R., Porcheron, A., Sweda, J.R., Jones, A.L., Mauger, E., & Morizot, F. (2016). Facial contrast is a cue for perceiving health from the face. *Journal of Experimental*

Psychology: Human Perception and Performance, 42, 1354-1362.

Saegusa, C., & Watanabe. K. (2016). Judgments of facial attractiveness as a combination of facial parts information over time: Social and aesthetic factors. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 173-179.

Said, C. P., & Todorov, A. (2011). A statistical model of facial attractiveness. *Psychological Science, 22*, 1183-1190.

Santos, I. M., & Young, A. W. (2011). Inferring social attributes from different face regions: Evidence for holistic processing. *Quarterly Journal of Experimental Psychology*, *64*, 751-766.

Scheib, J. E., Gangestad, S. W., & Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1431), 1913-1917.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8*, 350-353.

South Palomares, J. K., & Young, A. W. (2018). Facial first impressions of partner preference traits: Trustworthiness, status, and attractiveness. *Social Psychological and Personality Science, 9*, 990-1000.

South Palomares, J. K., & Young, A. W. (2019). Facial and self-report questionnaire measures capture different aspects of partner preferences. *British Journal of Psychology, 110*, 549-575.

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Burt, D. M., & Young, A. W.

(2013). Social inferences from faces: ambient images generate a three-dimensional model. *Cognition, 127*, 105-118.

Sutherland, C. A. M., Young, A. W., & Rhodes, G. (2017). Facial first impressions from another angle: how social judgments are influenced by changeable and invariant facial properties. *British Journal of Psychology, 108*, 397-415.

Thomas, D.R., Hughes, E., & Zumbo, B.D. (1998). On variable importance in linear regression. *Social Indicators Research, 45*, 253-275.

Todorov, A., Loehr, V., & Oosterhof, N.N. (2010). The obligatory nature of holistic processing of faces in social judgments. *Perception, 39*, 514-532.

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology, 66*, 519-545.

Todorov, A., & Porter, J. M. (2014). Misleading first impressions: Different for different facial images of the same person. *Psychological Science, 25*(7), 1404-1417.

Troje, N. F., & Bülthoff, H. H. (1998). How is bilateral symmetry of human faces used for recognition of novel views? *Vision Research, 38*, 79-89.

Thornhill, R., & Gangestad, S. W. (1993). Human facial beauty: Averageness, symmetry, and parasite resistance. *Human Nature, 4*, 237-269.

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*, 482-487.

Vernon, R. J. W., Sutherland, C. A. M., Young, A. W., & Hartley, T. (2014). Modeling first impressions from highly variable facial images. *Proceedings of the National Academy of Sciences, USA, 111*, E3353-E3361.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*, 592-598.

Young, A. W. (2018). Faces, people and the brain: the 45th Sir Frederic Bartlett Lecture. *Quarterly Journal of Experimental Psychology, 71*, 569-594.