Evolution and Human Behavior 38 (2017) 756–760

Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.ehbonline.org

Original Article

Women and men integrate facial information differently in appraising the beauty of a face

Daniel Gill *

Department of Statistics, The Hebrew University of Jerusalem, Israel
Institute of Neuroscience and Psychology, University of Glasgow, UK

ARTICLE INFO

Article history:
Initial receipt 11 July 2016
Final revision received 17 July 2017

Keywords:
Facial beauty
Perceptual integration
Analytical processing
Configural processing

ABSTRACT

Facial beauty plays a crucial role in social interactions, particularly in mating and reproduction. Therefore, the perceptual and cognitive mechanisms used for facial beauty assessment should be susceptible to different evolutionary and cultural pressures across genders and thus shape different observational appraising strategies. Using a novel approach, I evaluated the observers' subjective and unique importance given to specific facial attributes: eyes, nose, lips, and hair, and their spatial organization in the process of appraising the beauty of the whole face. These importance measures reveal the modulation of the integration of attributes strategy across the gender of observers and the sex of face. The degree of agreement about the beauty of the studied facial attributes was modulated across gender of observers and, for women observers, also across sex of face. Finally, I show that beauty appraisal can be mainly explained by a simple additive manner of isolated facial attributes appraisals.

1. Introduction

The beauty of faces is influential in many aspects of social interactions in general (Dion, Berscheid, & Walster, 1972; Little, Burriss, Jones, & Roberts, 2007) and in choice of mate in particular (Buss & Barnes, 1986; Walster, Aronson, Abrahams, & Rottman, 1966). Since the publication of Darwin's theory of natural selection (1859), the variability of perceived attractiveness has been analyzed in terms of the evolved signal content of striking phenotypic features, arguing that reproduction with a more attractive partner will increase an individual's biological fitness (Andersson, 1994; Barrett, Dunbar, & Lycett, 2002; Little, Burriss, Jones, DeBruine, & Caldwell, 2008). Choosing the right mate is crucial for successful reproduction, so reliable mechanisms for such recognition are favored by evolution. As a result, evolutionary, and maybe even cultural, pressures may act differently on women and men and, as a result, shape different observational beauty appraisal strategies across male and female genders.

In order to compare beauty appraisal strategies, one has to quantify the diagnostic dimensions of facial information that human observers use to judge the beauty of a face. Throughout history, several ideal characteristics of beauty have been suggested, mainly by formulating canons of face shapes and distances between selected facial landmarks of particularly meaningful and salient locations. The ancient Greeks believed aesthetic preferences fulfil certain geometrical conditions, such as the Golden Ratio. In the renaissance period, Neoclassical Canons were considered the ideal ratios of beautiful faces (Edler, 2001; Vegter & Hage, 2000).

Over the last few decades, many studies of facial beauty have focused on three main diagnostic dimensions: averageness, symmetry and sexual dimorphism (Gangestad, Thornhill, & Yeo, 1994; Langlois & Roggman, 1990; Perrett et al., 1998). On the other hand, the role of facial parts such as eyes, nose, and mouth, and their spatial organization and inter-attribute interactions (holistic processing) is a central issue in facial recognition research, suggesting different mechanisms and brain activation with single facial parts and their combinations (Arcurio, Gold, & James, 2012; Carey & Diamond, 1977; Farah, Wilson, Drain, & Tanaka, 1998; Gold, Mundy, & Tjan, 2012; Maurer, Grand, & Mondloch, 2002; Tanaka & Farah, 1993). The common view is that the human perceptual system integrates facial information into a gestalt whole rather than processing facial features in a non-interacting manner. The composite face effect has been used in many studies to demonstrate that facial parts cannot be perceived independently and therefore interact (Rossion, 2013; Young, Hellawell, & Hay, 1987). Nevertheless, there are some examples for which information conveyed from isolated facial parts is almost optimal when summed up in an additive manner (e.g., Maloney & Dal Martello, 2006). To date, the extent to which the impression of isolated facial parts shapes the assessment of facial beauty has not been studied.

What is the contribution of facial sub-regions and their spatial organization to the assessment of the beauty of the whole face? Pointing out the beauty of specific facial attributes is common in everyday life. The place of aesthetic characteristics of some facial attributes is well

* Department of Psychology, University of Winchester, Sparkford Road, Winchester SO22 4NR, UK.
E-mail address: Daniel.Gill@winchester.ac.uk.

http://dx.doi.org/10.1016/j.evolhumbehav.2017.07.001
1090-5138/

demonstrated by commonly used phrases, such as 'pretty eyes' or 'beautiful hair'. This suggests that facial beauty resides at different levels within the whole face at one level and at the level of 'facial parts' attributes at sub-levels. Nevertheless, the unique contribution of such specific sub-level attributes and the way they are integrated to make a beauty appraisal of the whole face, have not been investigated systematically and remain obscure.

Therefore, a prospective avenue for understanding the diagnostic dimensions which humans utilize to appraise facial beauty is an approach that rigorously quantifies the importance of the beauty of facial attributes, such as facial sub-regions and their spatial organization, to the beauty impression of the whole face.

Here, I address three questions about facial attributes processing for the purpose of beauty appraisal. Firstly, is the integration of facial attributes modulated by the gender of observer and the sex of face? Secondly, to what extent are the inter-subjective facial preferences modulated across facial attributes, gender of observer and sex of face? While observers may associate a similar degree of importance with certain facial attributes, they may disagree about the level of the beauty of individual attributes. A category of attributes which has a high level of agreement within a group of observers is an indication that there is a consensus, at least to some extent, about desirable specifications, such as shape or color, in that category. Such unique specifications may reflect a reliable signal of biological fitness or alternatively a social convention. Finally, to what extent is beauty appraisal based on the additive processing of facial attributes?

In the current study, I quantitatively evaluate the unique contribution of specific facial attributes to the beauty appraisal of whole faces. I use these measures to investigate how the integration strategy is modulated across the gender of observers and across the sex of face. Later, I study the modulations of inter-subjective homogeneity across the gender of observers and across the sex of face. Finally, I show that the majority of the feasible variance of beauty appraisal of the whole face is explained by the appraisal of the isolated attributes I used in the current study.

The facial phenotype is derived by the biological sex; therefore throughout this paper, I classify the face stimuli by their biological sex: female or male (Enlow, 1996). However, since it is unknown which factors shape the strategy of beauty perception, biological or cultural; I have chosen to follow the common distinction used in cross-gender studies and classify the observers by the term 'gender': women or men.

2. Method

2.1. Observers

Sixty four observers (32 women, M = 22.8, SD = 2.3 years; 32 men, M = 23.8, SD = 2.7 years) participated in a task rating the female face. Sixty four observers (32 women, M = 22.4 years, SD = 1.9 years; 32 men, M = 23.8 years, SD = 3.2 years) participated in a task rating the male face. This sample size was determined in advance. As a data driven study utilizing a novel method, the types of effects and their expected sizes were unknown. All observers were students at the Hebrew University of Jerusalem, with normal or corrected to normal visual acuity, who participated in the experiment for course credit or monetary reward. All observers signed an informed written consent according to the institutional review board of the Hebrew University of Jerusalem.

2.2. Stimuli and apparatus

I used two sets of frontal headshot color photographs of individuals with neutral expressions: one set of 27 Caucasian females and one set of 27 Caucasian males (all models aged between 20 and 30). The faces had similar location, size, illumination, and there were no beards, moustaches, earrings, eyeglasses, makeup, or jewellery. The resolution of all images was 350 × 480 pixels and the models had been instructed to assume neutral expressions. Four facial fragments were cut out from the intact faces: eyes (including eyebrows), nose, mouth, and hair (including ears, seen or occluded). An additional stimulus category denoted here as 'configuration', was made to capture the spatial organization of the eyes, nose and lips together with facial shape elements. I denote the latter category as 'configuration', however this should not be confused with the identically named term sometimes used in other studies. To create the configuration stimuli, images of the whole face were converted into greyscale (to partial out the facial coloration contribution leaving only the luminance channel), then low-pass filtered with a critical band of approximately six cycles per face width (to partial out the inner facial features specification; see Goffaux, Hault, Michel, Vuong, & Rossion, 2005), and finally cropped of hair. Fig. 1 illustrates the six categories of stimuli: eyes, nose, lips, hair, configuration, and the whole face. All stimuli were presented on a 17 inch LCD screen at a viewing distance of 60 cm.

2.3. Procedure

Each observer participated in six different conditions, each focusing on a different category: eyes, nose, lips, hair, configuration, and whole face. The first five conditions were blocked by attribute and presented in random order of blocks and random order of individual stimuli within blocks across participants. The whole face condition was always presented as the final block in a random order of stimuli within blocks.

In each condition, pairs of images (of the same attribute and sex of face, e.g., two pairs of male noses) were presented on screen, side by side, in a random order and a random left/right juxtaposition. Participants were instructed to indicate, using a five alternative forced choice method, which of the two images they thought was more beautiful: 'the left image is much more beautiful'; 'the left image is slightly more beautiful'; 'both images are equally beautiful'; 'the right image is slightly more beautiful', and 'the right image is much more beautiful'. In most studies that address the aesthetic aspects of faces and body the term 'attractiveness' is typically used. Nevertheless, in the current study the participants were instructed to indicate the 'beauty' and not the 'attractiveness' of the face as the latter term can be interpreted also in terms of sociability and may lead to different interpretations across participants (e.g., in the case of a 'mean but beautiful' face).

3. Results

The beauty score of an individual stimulus was derived from the pairwise comparison in the following way. For each trial, if an individual stimulus was rated in a single pairwise comparison as 'much more beautiful' than the other, it got the value 2 and the other, less beautiful individual stimulus, got the value −2. In a similar way, the 'more beautiful' response yielded ratings of 1 to the more beautiful stimulus and −1 to the less beautiful stimulus. 'Equally beautiful' was evaluated as 0 for both stimuli. Fig. 2A illustrates the data pre-processing stage: to obtain a unique subjective score for each individual stimuli and each observer, I averaged the responses for each observer over all comparisons in which the individual stimuli took part. To avoid heterogeneity in the use of the response scale among participants and stimulus categories, the average responses were converted to ranks over identities within each subject and each category of stimulus. This pre-processing step yielded a subjective beauty score for each individual stimulus and each observer. To measure the importance of each facial attribute to the whole face, I used the semipartial correlation between each of the attribute scores and the matching scores of the whole face (Darlington, 1990). This statistic provides some desirable properties: (i) the semipartial correlation measures the exclusive contribution of the attribute in question to the whole face appraisal whilst partialing out the rest of the facial attributes from that attribute, in other words, it measures the contribution of the specific attribute to the whole face appraisal that cannot be explained by any of the other attributes; (ii)