Within- and Between-Category Face Identity Discrimination: Association with Physical Facial Features

Tor Ekstrom[1], Stephen Maher[1], Daniel Norton[2] and Yue Chen[1,*]

[1]*McLean Hospital, Department of Psychiatry, Harvard Medical School*

[2]*Department of Psychology, Boston University*

Abstract: Facial identity discrimination is crucial for success in the social world. Yet, how this task is accomplished remains incompletely understood. Face perception involves identifying members of same or different of social categories (such as gender and race). We know that differences in facial morphology are critical for discriminating individual faces. What is unclear, however, is whether the same mechanisms are used to discriminate identities when they are members of the same versus different social categories. To address this question, this study examined and compared perceptual performance on within- and between-category facial identity discrimination tasks and evaluated the relationship between the perceptual performances and physical features of face images used for discrimination. Face images from each of five pairs of individuals (two Asian females, two Caucasian females, two African males, one African male and one Caucasian male, or one Caucasian female and one Caucasian male) were morphed to create images along a continuum of facial dissimilarity. Using the original and created images, with psychophysical methods we measured perceptual thresholds for facial identity discrimination in a group of human observers (n=24). For each pair of original face images, the differences in six physical facial features (end-lip raise, mid-top-lip raise, mid-top-lip, mid-low-lip, eye-opening, mid-eyebrow raise and luminance) were also measured. Perceptual thresholds were the lowest (best performance) for the across-race condition, the second lowest for the across-gender condition and the highest for the three same-race and same-gender conditions. Two physical facial features, mid-top-lip raise and luminance, each accounted for a significant portion of perceptual performance (35% and 40%, respectively). These results indicate that between-category facial identity discrimination is more precise than within-category discrimination. Precise discrimination of facial identity is associated with selective physical facial features.

INTRODUCTION

Discriminating the identity of faces plays a crucial role in social life. This cognitive ability is in principle limited by the physical differences of faces being discriminated, for example, the varying distances between two individual's eyes. A prominent model of face perception holds that information from a face is coded relative to a hypothetical average face at the center of a mental face-space [1, 2]. According to this model, perceptual discrimination of two faces is constrained by the distance between the faces in face-space, that is, by the dissimilarity in their physical appearance. Individuals who belong to different social categories (such as race and gender) usually have distinct physical facial features [3, 4]. It is thus reasonable to assume that facial identity discrimination is easier for between-category conditions than for within-category ones. These assumptions have rarely been tested empirically however, and we have little exact knowledge about the differential perceptual capacities for within- and between-category facial identity discrimination. Although it has been shown that eye

height and mouth height provide important cues for facial identity discrimination [5, 6], our overall knowledge about specific facial features that support facial identity discrimination is limited.

In this study, we evaluated two specific questions: 1) if perceptual discrimination of facial identities is affected by the social categories that those identities are associated with, and 2) if differences in facial identity discrimination of same vs. different social categories can be accounted for by physical feature differences between the faces being discriminated.

To compare perceptual performances within same category with those among different categories, we used three sets of conditions: 1) across-gender (Caucasian female vs. Caucasian male), 2) across-race (African male *vs.* Caucasian male), and 3) same-race and same gender (Asian females, Caucasian females or African males). To determine the precision of face identity discrimination, we systematically manipulated facial dissimilarity levels under each condition through morphing between sets of two original facial images and then measured corresponding perceptual thresholds in a group of human observers.

To compare perceptual performance with the change in physical features drawn from face images being

*Address correspondence to this author at the MS 303, 115 Mill St. Belmont, MA 02478; Tel: 1 617 855 3615; Fax: 1 617 855 3611; E-mail: ychen@mclean.harvard.edu

discriminated, we measured six physical differences between the facial images used for each condition (end-lip raise, mid-top-lip raise, mid-top-lip, mid-low-lip, eye-opening, mid-eyebrow raise and luminance) [7]. Correlations between perceptual performances of human observers and change in physical features in face images were then computed.

Our working hypothesis was that facial identity discrimination is more precise for the between-category conditions than for the within-category conditions. Perceptual performance is correlated with change in physical features between the face images being discriminated; the greater the facial feature differences, the more precise facial identity discrimination.
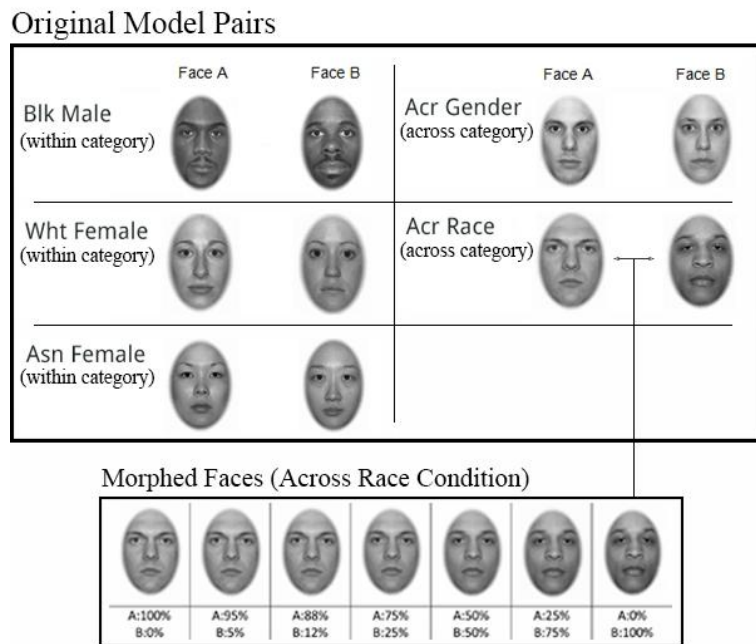
METHODS

Stimuli

Visual stimuli were images of ten face models taken from the NimStim Face Stimulus Set (http://www. macbrain.org) [8]. All face images used here had neutral expressions. The face images were cropped to exclude hair, clothing and other visual cues, using Adobe Photoshop (CS6, version 13). The processed images contained only a face, and were resized to fit a rectangular area of 265 x 355 pixels (subtending 13 x 17 degrees of visual angle). They were paired according to two categories - race and gender. The selected face pairs were 1) Caucasian females: Model #2 and #6. 2) African males: Model #39 and #43, 3) Asian females: Model #15 and #16. 4) Across-race: Model #26 (Caucasian male) and #42 (African male), and 5) Across-gender: Model #8 (Caucasian female) and #34 (Caucasian male) (Fig. **1**).

Additional face images were created by morphing (FantaMorphPro 5.0, 2012) between two original images (A and B) for each pair of face model. There were five created levels of facial difference for each pair: 5% (in which 5% of Face A and 95% of Face B were combined to form a created face image), 12%, 24%, 48%, and 75% (Fig. **1**). In combination with the two original face images of each pair: 0% (un-morphed Face B) and 100% (un-morphed Face A), seven levels of facial difference were used for this study. Stimulus presentation, as well as the recording of subjects' responses, were programmed within Psychtoolbox 3 [9].
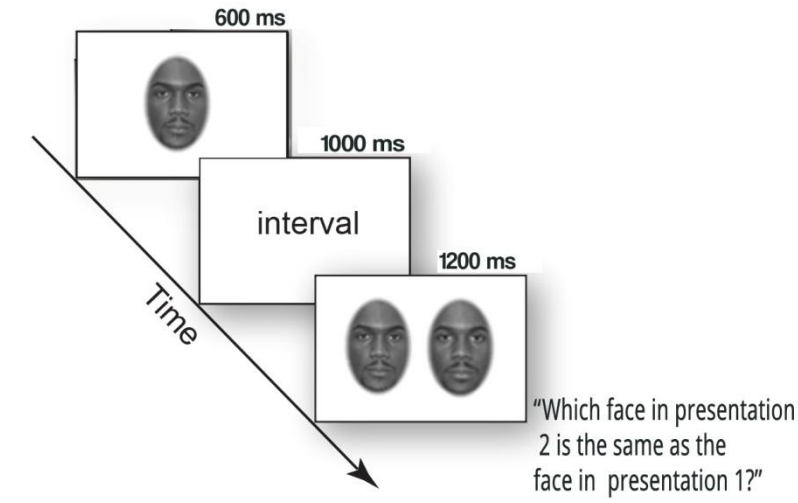
Design and Procedure

Stimuli were presented using a method of constant stimuli. Each trial included two presentations: First, a single face was displayed for 600 msec. Second, a pair of faces was presented for 1200 msec. There was a 1000 msec interval between two presentations. Participants' task was to indicate which of the two faces in the second presentation (left or right) was the same as the face in the first presentation (Fig. **2**). They performed this two-alternative forced choice (AFC) task by pressing the left-arrow key to indicate if the left face in the second presentation matched the face in the first presentation, and the right-arrow key to indicate if the right face in the second presentation matched the face in the first presentation.

We used a design of 5 face model pairs by 7 facial difference levels by 2 locations (left or right) by 8 repeats for the second presentation. The trials were first blocked by the five face model pairs. Within each block, the trials of different facial difference levels and locations were

Fig. (1). Illustration of original and morphed face stimuli.
Samples of original (un-morphed) and created face images used in this study. The top panels listed five pairs of original images. The bottom panel included the created images for one pair of original images (Across race, Face A and Face B). The created images were produced through morphing between the two original images, and contained various percentages of Face A (0, 5, 12, 25, 50, 75, 88, 95, or 100%) and Face B (100, 95, 88, 75, 50, 25, 12, 5, or 0%).

Fig. (2). Illustration of the facial identity discrimination paradigm.

Schematic illustration of study procedure. In the first presentation of a trial, an original face image was shown. In the second presentation, two face images, one original and one created, were shown side by side. Subjects' task was to indicate which of the two images in the second presentation, left or right, was the same as the one in the first presentation.

counterbalanced. Note that the trials for 0% facial difference level included only one location (two identical images). Each participant performed 570 trials in total.

The main performance measure was the percent of correct responses, or accuracy, under each condition. Reaction time (RT) for each condition was also recorded. In addition, perceptual threshold of facial identity discrimination, defined as the minimum facial difference level to yield 80% accuracy, was derived as a separate performance measure for each subject [10, 11].

Two additional cognitive measurements, verbal IQ (not using face-related information) and theory of mind (using face-related information), were obtained to compare with facial identity discrimination. For each subject, verbal IQ was measured using the verbal component of the WAIS-R [12]. Theory of mind was measured using the Eyes Test [13].

Subjects

Twenty-four human observers (8 females) participated in the study. Participants were between 18 to 65 years old (mean age = 36.91, SD= 11.92). Racially, three participants identified as African American, two as Asian American, sixteen as Caucasian, and three participants declined to respond. Ethnically three participants identified as Hispanic or Latino. All participants were screened for exclusion of neurological or psychiatric conditions or recent history of substance abuse (past 6 months). Participants had normal or corrected to normal vision, as indexed using the Rosenbaum vision screening chart. The subjects were recruited from the local community and gave written informed consent prior to participation. The research protocol was approved by the Institutional Review Board of McLean Hospital.

RESULTS

A two-way ANOVA (5 face model pairs x 7 facial difference levels) for accuracy showed significant main

effects of face model pair (df=4, F=5.84, p<0.001), and facial difference level (df=6, F=292.09, p<0.001) (Fig. **3** top panel). The interaction between face model and facial difference was not significant (df=24, F=1.02, p=0.44). This indicates that perceptual performances differed for these face model pairs, and were improved with the increase of facial difference levels. A pairwise analysis comparing the relationship of perceptual thresholds between every face model pair yielded a significant difference only between the across-race and the rest of face model pairs (df=1, F=18.85, p<0.001), indicating performance was significantly better for the across-race face model pair.

A two-way ANOVA (5 face model pairs x 7 facial difference levels) for RT showed significant main effects of face model (df=4, F=3.78, p=0.005) and facial difference (df=6, F=11.26, p<0.001), but no interaction (df=24, F=0.1, p=1.0). This again indicates that reaction times differed for these face model pairs, and decreased with the increase of facial difference levels. The only significant pairwise comparison for face model was between the across-race and the within-category model pairs (df=1, F=5.82, p=0.016), indicating that reaction time was significantly shorter for the across-race than for the within category face model pairs.

One-way ANOVA (5 face model pairs) of perceptual threshold showed a trend towards a significant main effect (df=1, F=2.00, p=0.11). However, when comparing the within-race and within-gender pairs with the across-race face model pairs, a significant difference was found (df=1, F=8.234, p=0.005) (Fig. **3** bottom panel). While also low, perceptual thresholds for the across-gender condition were not significantly different when compared with within-category conditions (df=1, F=1.349, p=0.249).

Among the five face model pairs tested, the across-race pair yielded the largest percent changes in four out of the seven facial features measured. Detailed percent changes of

all facial features for all 5 face model pairs are listed in Table **1**.

The correlations between perceptual thresholds of facial identity discrimination and facial feature changes are listed in Table **2**. Substantial inverse correlations were found between perceptual thresholds and mid-top-lip change, and between perceptual thresholds and luminance change. That is, the greater the changes in these two facial features, the lower the threshold (i.e. the better the perceptual performance) (Fig. **4**). Here the perceptual thresholds were

A.

Identity Discrimination

B.
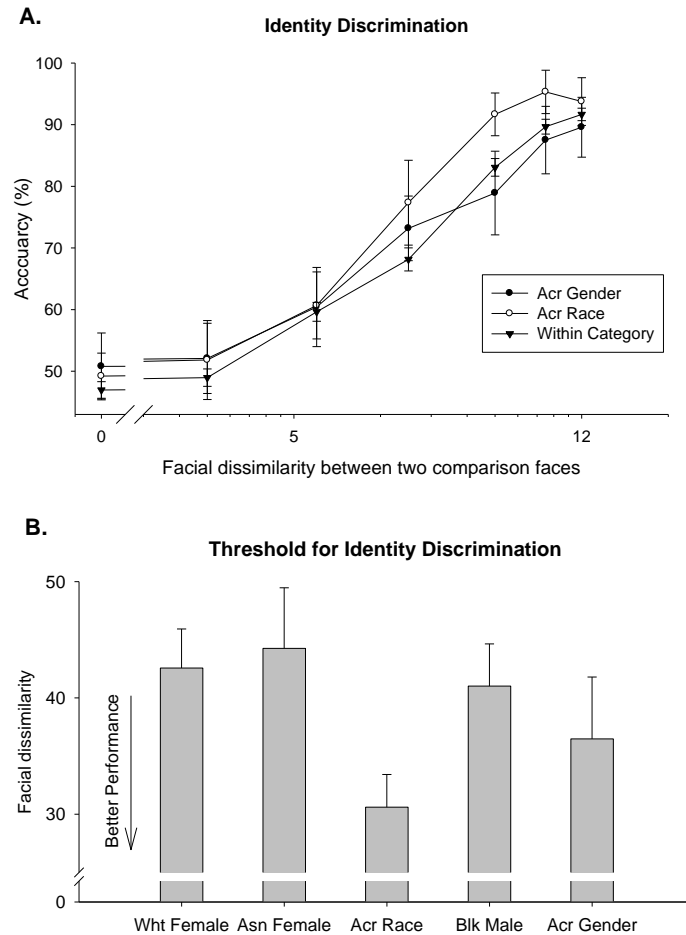
Threshold for Identity Discrimination

Fig. (3). Summary of performance on facial identity discrimination.

Perceptual accuracy and threshold of face identity discrimination. In the top panel (**A**.), x axis represents the dissimilarity level of facial images used for perceptual discrimination (in logarithmic scale). Y axis represents the percents of correct trials, or the accuracy, during the performance of the task. Error bars indicate ±1 standard error. In the bottom panel (**B**.), x axis represents the category to which two original face images are belong (e.g., across race). Y axis represents the perceptual threshold during the performance of the task, which was defined as the minimum dissimilarity level of facial images under which the 80% correct of performance level was reached, and was derived from the best fitting psychometric function (Weibull) to observed accuracy data for each subject In both panels, error bars indicates ±1 standard error.

Table 1.   Facial feature changes between the two un-morphed images of each face model pair.

Facial feature	Caucasian Female	Across Race	Across Gender	Asian Female	African Male
end-lip raise (line1)	2.46	3.32	4.29	2.18	4.39
mid-top-lip raise (line 4)	6.45	3.64	4.39	7.10	7.29
mid-top-lip (line 6)	22.17	56.89	31.61	1.11	37.43
mid-low-lip (line 7)	7.47	25.93	13.27	14.19	16.71
eye-opening (line 8)	5.29	39.35	11.72	19.64	13.59
mid-eyebrow raise (line 12)	14.33	4.51	2.33	8.16	1.61
Luminance	5.52	16.97	1.16	1.45	2.07

averaged across all subjects tested. Thus correlation values were derived between 5 average thresholds and 5 feature changes from each face pair. To see whether these averaged-threshold base correlations were significantly different across these 5 pairs of faces, individual perceptual thresholds (n=24 for each of 5 pairs) were used. The individual threshold based correlations were indeed significantly different from each other. The set of Pearson correlations between line 6 and lines 1 and 4 were significantly different (p≤0.001). The correlation between line 6 and line 12 also approaches significant difference (p=0.087). This analysis of individual thresholds indicates that mid-top-lip (line 6) is preferentially related to perceptual performance.

To compare with non face variables, perceptual thresholds for five pairs of face images were averaged for each subject. None of the correlations were significant between the averaged perceptual thresholds for facial identity discrimination and Verbal IQ (n=22, ps<.05) or Eye Tests scores (a social cognitive index) (n=19, ps>.05) (Table **3**). The perceptual performances were also not significantly correlated with verbal IQ scores (a general intelligence index) (Table **3**).
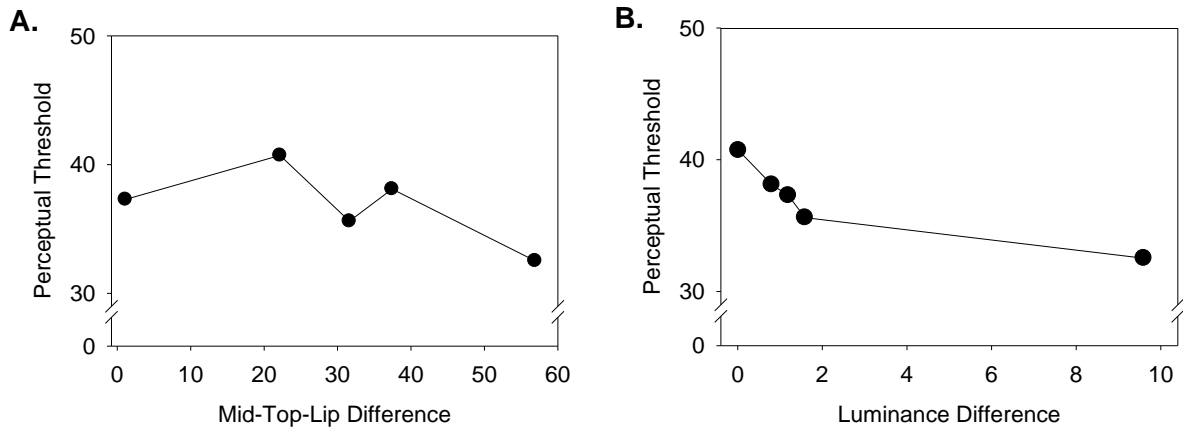
DISCUSSION

This study shows that perceptual discrimination of facial identities is more sensitive for the across race category than same race and gender categories. This study also shows that the sensitivity of facial identity discrimination is associated with two physical facial features (luminance and mid-top-lip). These results suggest that perceptual discrimination of facial identity is linked to selected facial features.

Perceptual Discrimination of Facial Identity

It has been assumed, but rarely assessed, that perceptual discrimination of face identity depends upon physical differences between the faces being discriminated. Using psychophysical methods, this study showed that to reliably discriminate the identities of two individuals (i.e. at 80% accuracy level), about 30 to 40 percent (Fig. **2**) of overall physical difference between face images is required. The greater this physical difference between two individual face images, the more precise facial identity discrimination will be. This result has two implications. First, it confirms that facial identity discrimination can be understood under the face space model, in which the distance between two faces in the mental face space is a key factor to discriminating the faces. In general, the farther apart the two faces are in the mental face space, the more easily they can be differentiated from each other. Specifically, performance levels of perceptual discrimination are proportional to the difference in selected, not overall, facial features (such as mid-top-lip).

Fig. (4). Relationship between perceptual performance and change in the two physical facial features.
Relationship between perceptual thresholds and changes in physical facial feature (Panel A: mid-top-lip and Panel B: luminance). Perceptual thresholds were averaged across subjects. Changes in the two physical facial features were measured between the two original (un-morphed) face images in each face model pair.

Table 2. Correlation coefficients between perceptual threshold and change in physical facial feature.

Facial feature	Pearson's R
end-lip raise (line1)	0.020
mid-top-lip raise (line 4)	0.241
mid-top-lip (line 6)	-0.345
mid-low-lip (line 7)	-0.265
eye-opening (line 8)	-0.280
mid-eyebrow raise (line 12)	-0.050
luminance	-0.395

Table 3.   Pearson's R for correlation between perceptual thresholds and Verbal IQ and Eyes Test scores.

Face Model Pair	Verbal IQ v. Threshold	Eyes Test v. Threshold
White Female	0.012	0.402
Asian Female	-0.236	0.128
Acr Race	-0.072	-0.025
Black Male	-0.208	-0.132
Acr Gender	0.063	0.093
Average	-0.322	-0.063

Second, the use of only 30 to 40 % of facial differences suggests that our face processing system has a redundant capacity for distinguishing different facial identities at perceptual levels. More often than not, face images encountered in the social world may be degraded, which presents a challenge for perceptual processing of facial images [14]. Redundant capacity of facial identity discrimination would allow sufficient perceptual processing of degraded facial signals for this socially important function.

Discrimination of Face Identity in the Context of Social Category

Greater physical differences of certain facial features belonging to different social categories presumably make it easy to distinguish identities associated with those faces. This study supports this presumption. The highest and the second highest performance levels were achieved for the across-race and the across-gender face model pairs, with lower performance levels for the three same-race and same-gender face model pairs. It has been suggested that facial judgment of social categories is optimal through category-specific facial features [15]. Our result of the more precise perceptual discrimination for the across-race category than for the within-category conditions suggests that the categorization of race, formulated at social construct levels, has its root at perceptual processing levels, as far as facial identity is concerned. Recently, it has been shown that fusiform face area (FFA), a cortical region responsible for perceptual processing of facial information [16-18], also subserves the neural processing of face-based race and gender information [19-21]. The results of these neuroimaging studies and the present study are all consistent with the notion that perceptual and social category information from faces is at least partially processed jointly in the brain.

Physical Facial Features for Facial Identity Discrimination

As expected, the differences in selected facial features between the two face images of each face model pair were larger for the between than for the within-category conditions. Perceptual discrimination, also as expected, was more precise for the between than for the within-category conditions. For example, the highest perceptual performance levels were achieved for the across-race face model pair which also corresponds to the largest facial feature difference (mid-top-lip, eye-opening and luminance) (Table **1**). This link between the two sets of variables suggests that physical features of face images play a crucial role in facial identity discrimination. Note that it was not any particular feature that can singularly account for perceptual performance. For example, there was a substantial change in mid-top-lip for the African male face model pair (37.43%, the second largest among the five face model pairs) (Table **1**), yet the corresponding perceptual performance was not as precise (Fig. **2**). This suggests that it is a combination of selected facial features that collectively play a crucial role in perceptual discrimination of facial identity.

Using a psychophysical approach, this study produces supporting evidence for the two general assumptions on facial identity discrimination -1) better perceptual performance under the across-race than the within-category conditions and 2) association of physical facial features and perceptual performance. Two caveats of this study however need to be noted. First, the results here were obtained with a set of face images that is modest in size for a systematic comparison of between-category and within-category perceptual discrimination. Gender and race of observers may affect the performance levels when face stimuli possess different genders and races [22-24]. In future studies, a larger set of face images may afford to disassociate this factor from perceptual discrimination. Second, the luminance and contrast were not systematically equalized across the face images used here. While a physical facial feature, luminance values are not directly related to facial structure per se. It is unclear how perceptual performances for the between- and within-category facial identity discrimination are affected when this general visual variable is removed. It is unlikely however, that luminance played a major role in perceptual performance here. While both the highest perceptual performance results and the greatest luminance difference occurred under the across-race condition, the next two highest performance results (across-gender and black male pairs respectively) did not occur to the next two greatest luminance differences in face images (Table **1**). Application of a larger set of contrast and luminance-equalized face images in future studies would be crucial for further

characterizing the perceptual capacities of discriminating faces of the same and different categories and the physical constraints of such capacities.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Valentine T. A unified account of the effects of distinctiveness, inversion, and race in face recognition. Q J Exp Psychol A 1991; 43(2): 161-204.

[2]   Lee K, Byatt G, G. Rhodes. Caricature effects, distinctiveness, and identification: testing the face-space framework. Psychol Sci 2000; 11(5): 379-85.

[3]   Ferrario VF. Sexual dimorphism in the human face assessed by euclidean distance matrix analysis. J Anat 1993; 183 (Pt 3): 593-600.

[4]   Farkas LG. International anthropometric study of facial morphology in various ethnic groups/races. J Craniofac Surg 2005; 16(4): 615-46.

[5]   Susilo T, McKone E, Edwards M. What shape are the neural response functions underlying opponent coding in face space? A psychophysical investigation. Vision Res 2010; 50(3): 300-14.

[6]   Robbins R, McKone E, Edwards M. Aftereffects for face attributes with different natural variability: adapter position effects and neural models. J Exp Psychol Hum Percept Perform 2007; 33(3): 570-92.

[7]   Pilowsky I, Katsikitis M. The classification of facial emotions: a computer-based taxonomic approach. J Affect Disord 1994; 30(1): 61-71.

[8]   Tottenham N. The NimStim set of facial expressions: judgments from untrained research participants. Psychiatry Res 2009; 168(3): 242-9.

[9]   Brainard DH. The Psychophysics Toolbox. Spat Vis 1997; 10(4): 433-6.

[10]   Chen Y. Visual and cognitive processing of face information in schizophrenia: detection, discrimination and working memory. Schizophr Res 2009; 107(1): 92-8.

[11]   Norton D. Association of impaired facial affect recognition with basic facial and visual processing deficits in schizophrenia. Biol Psychiatry 2009; 65(12): 1094-8.

[12]   Wechsler D. Manual for the adult intelligence scale-revised. Psychological Corporation: New York, 1981.

[13]   Baron-Cohen S. The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. J Child Psychol Psychiatry 2001; 42(2): 241-51.

[14]   Gilad-Gutnick S, Sinha P. Recognizing degraded faces: the contribution of configural and featural cues. Perception 2012; 41(12): 1497-511.

[15]   Cloutier J, Mason MF, Macrae CN. The perceptual determinants of person construal: reopening the social-cognitive toolbox. J Pers Soc Psychol, 2005; 88(6): 885-94.

[16]   Kanwisher N, Tong F, Nakayama K. The effect of face inversion on the human fusiform face area. Cognition 1998; 68: p. B1-11.

[17]   Kanwisher N, McDermott J, Chun MM. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci 1997; 17(11): 4302-11.

[18]   Haxby JV, Hoffman EA, Gobbini MI. Human neural systems for face recognition and social communication. Biol Psychiatry 2002; 51(1): 59-67.

[19]   Ratner KG, Kaul C, Van Bavel JJ. Is race erased? Decoding race from patterns of neural activity when skin color is not diagnostic of group boundaries. Soc Cogn Affect Neurosci 2012. 8(7): 750-5.

[20]   Kaul C, Rees G, Ishai A. The gender of face stimuli is represented in multiple regions in the human brain. Front Hum Neurosci 2011; 4: 238.

[21]   Contreras JM, Banaji MR, Mitchell JP. Multivoxel patterns in fusiform face area differentiate faces by sex and race. PLoS One 2013; 8(7): e69684.

[22]   Heron-Delaney M. Perceptual training prevents the emergence of the other race effect during infancy. PLoS One 2011; 6(5): e19858.

[23]   McBain R, Norton D, Chen Y. Females excel at basic face perception. Acta Psychol (Amst) 2009; 130(2): 168-73.

[24]   McBain R, Norton, and Y. Chen, A female advantage in basic face recognition is absent in schizophrenia. Psychiatry Res 2010; 177(1-2): 12-7.