[Accepted]

Facial Structure and Perception of Sexual Orientation: Research with Face Models Based on Photographs of Real People.

ABSTRACT

Some evidence suggests that lay persons are able to perceive sexual orientation from face stimuli above the chance level. A morphometric study of 390 heterosexual and homosexual Canadian people of both sexes reported that facial structure differed depending on the sexual orientation [Skorska et al., 2015]. Gay and heterosexual men differed on three metrics as the most robust multivariate predictors, and lesbian and heterosexual women differed on four metrics.

A later study [xxxxxxxxxxxx] verified the perceptual validity of these multivariate predictors using artificial 3D face models created by manipulating the key parameters. Nevertheless, there is evidence of important processing differences between the perception of real faces and the perception of artificial computer-generated faces.

The present study, composed of two experiments, tested the robustness of the previous findings and extended the research by experimentally manipulating the facial features in face models created from photographs of real people. Participants of the Experiment 1 achieved an overall accuracy (0.67) significantly above the chance level (0.50) in a binary hetero/homosexual judgment task, with some important differences between male and female judgments. On the other hand, results of the Experiment 2 showed that participants rated the apparent sexual orientation of series of face models created from natural photographs as a continuous linear function of the multivariate predictors. Theoretical implications are discussed.

KEYWORDS: Sexual orientation, Facial structure, Face Models, Photographs, Homosexuality, Perception.

When two persons meet for the first time, certain individual characteristics, such as age, sex, and race, are easily inferred at first sight, but other features, such as sexual orientation, are not perceptually obvious. An issue examined in research on personal and social perception has to do with the common belief that sexual orientation can be spontaneously perceived on the basis of certain nonverbal information (Rule, 2016). Unfortunately, many homosexual people in democratic countries still experience subtle discrimination due to their sexual orientation. Furthermore, many countries around the world do not recognize the rights of LGBT (lesbian, gay, bisexual, and transgender) people, and most of them are victims of sexual prejudice and persecution.

Beyond the social implications, the possible perception of a person's sexual orientation at first sight is a question that has received attention in psychological research. A meta-analysis performed by Tskhay and Rule (2013) found a significant correlation of $r = .29$ between actual (reported by the individuals themselves) and perceived sexual orientation by part of experimental participants. This review showed that participants could recognize sexual orientation above the chance level observing different types of stimuli, such as photographs (Rule & Ambady, 2008) or video clips (Valentova, Rieger, Havlicek, Linsenmeier, & Bailey, 2011. Nevertheless, some evidence suggests that heterosexual individuals tend to look, express emotions, and act more typical for their sex compared to homosexual individuals, and people can use this information based on social stereotypes to infer sexual orientation (Bjornsdottir, R. T., & Rule, N. O. (2020). Consequently, it would be quite encouraging to find some cues informative of sexual orientation devoid of any social or cultural influence, and anatomical facial structure *per se* could be a good candidate.

Human faces are a significant source of individual information, and some studies reveal that a certain degree of success can be obtained in judgments about sexual orientation from a whole face (e.g., Freeman, Johnson, Ambady, & Rule, 2010; Rule & Ambady, 2008) or from isolate facial parts (Rule, Ambady, Adams, & Macrae, 2008). In addition, experimental results suggest that such judgments are carried out unconsciously (Rule, Macrae, & Ambady, 2009). Although these perceptions can be based on noticeable cultural cues, such as makeup, earrings, haircut, lipstick, and so forth, experimental data suggest that accurate judgments based on subtle and non-obvious anatomical information are possible (Rule et al., 2008).

In recent years, a few works have studied the possible objective relationship between sexual orientation and anatomy of the facial structure. Hughes and Bremme (2011) found that photographs from self-identified heterosexual persons (in webpages) displayed less-asymmetrical facial measures than self-identified homosexual persons, but the two groups did not differ significantly on other facial metrics. Valentova, Kleisner, Havlicek, and Neustupa (2014) performed a morphometric analysis of facial photographs of eighty heterosexual and homosexual men from the Czech Republic, and they obtained that, on average, the two groups differed in a set of facial features. Homosexual men presented relatively wider and shorter faces, more rounded jaws, and smaller and shorter noses. Interestingly, the authors found that differences in facial morphology of heterosexual vs. homosexual men did not merely mirror variation in masculinity- femininity rates made by independent raters.

More recently, it is necessary to highlight the massive morphometric study that Skorska, Geniole, Vrysen, McCormick, and Bogaert (2015) performed in Canada on almost four hundred homosexual/heterosexual men and women. After inputting a facial photograph of each person into FaceGen (Singular Inversions, 2010), a software that provides 62 different facial metrics, the authors performed statistical analyses as a function of the sexual orientation. At a first univariate analysis, homosexual (lesbian) and heterosexual women on average yielded significant differences on a set of 17 facial measures, whereas men showed differences on a set of 11 parameters. In a second step, the authors carried out a multivariate analysis and could isolate four important predictors of sexual orientation in women and three in men. Specifically, gay men had shorter noses, more convex cheeks, and foreheads that tilted back more. On average, lesbian women had more turned up noses, mouths that were more puckered (less retracted), smaller foreheads, and marginally more masculine face shapes. The authors discussed possible explanations, some related to prenatal and postnatal sex hormone exposure.

A question arising from Skorska et al.'s (2015) morphometric study is whether the main facial characteristics identified could have any effect on the perception of sexual orientation by part of lay people. This hypothesis was recently tested by González-Álvarez (2017) by experimentally manipulating these facial parameters in artificial 3D faces created by computer. The study consisted of two experiments in which the artificial stimuli were presented to Spanish young adults (mean age around 21 years old) of both sexes (more female participants). The first experiment involved a binary hetero/homosexual judgment task that obtained a significant accuracy of 74%. In the second experiment, five versions of twenty-four face models of both sexes were created, modifying the facial attributes in equal steps. The task consisted of judging on a scale the apparent degree of homosexuality of each face, and the rating scores almost perfectly fit a linear regression model.

In summary, the experiments by González-Álvarez (2017) supported the perceptual validity of the facial predictors identified by Skorska et al. (2005) for both men and women. It should also be noted that the sample of participants belonged to an environment (Spain) culturally and geographically different from Canada. However, the González-Álvarez's study utilized totally artificial faces created *ex novo* by computer. In the scientific literature, some evidence suggests important differences in the processing of natural versus artificial faces (Balas & Pacella, 2015; Carlson, Gronlund, Weatherford, & Carlson, 2012; Kätsyri & Sams, 2008; see also González-Álvarez & Cervera-Crespo, 2019) and also differences in the neural representation of both types of faces (Wheatley, Weinberg, Looser, Moran, & Hajcak, 2011). It seems that artificial faces are more difficult to recognize and remember than natural faces, and some emotions elicited by artificial vs. natural faces are processed differently (Ehrlich, Schiano, & Sheridan, 2000; Kätsyri & Sams, 2008). Moreover, current psychological science (especially social psychology) is experiencing a certain replication crisis (Open Science Collaboration, 2015; see also recently Baucal, Gillespie, Krstić, & Zittoun, 2020), and there is some concern about the reproducibility and robustness of experimental results, particularly in social psychology. The aim of the present study was to test the robustness of the observations of González-Álvarez (2017) and to extent the research by using another type of stimuli (face models created from photographs of real people) that is presumably processed differently.

EXPERIMENT 1

The purpose of this experiment was to test whether lay people can discriminate between two versions of face models based on photographs of real people, which (the versions) differed only in the facial attributes found by Skorska et al. (2015) as multivariate predictors of sexual orientation. We hypothesized that participants will be able to choose the "homosexual" version of each realistic face model significantly above the chance level (0.50).

METHOD

Participants

Participants were seventy-six adults of both sexes (62 females), whose age range was 18-37 years (M = 19.34; SD = 2.64). All of them were undergraduates at the University xxxxxxxxxxxxxxx, who voluntarily participated in the experiment as course credit. Of those who indicated ethnicity (n=72), 75% were White/Caucasian, 19% were Hispanic/Latin American, and 6% were another ethnicity.

Materials

The stimuli consisted of 96 realistic 3D face models created with the FaceGen Modeller 3.5 software (Singular Inversions, 2010) as follows. First of all, 48 “neutral” face models (not used in the experiment) were generated from photographs of real people (24 men and 24 women, all Caucasian except two people) selected from two international face databases. In concrete, the face datasets were the Basel Face Database (Walker, Schönborn, Greifeneder, & Vetter, 2018) from the University of Basel, Switzerland, (photographs of individuals 1, 3, 4, 5, 11, 20, 21, 27, 32, 33, 34, 40 for men, and 2, 6, 10, 12, 15, 18, 19, 22, 26, 35, 37, 38 for women were selected) and the Karolinska Directed Emotional Faces (KDEF) (Lundqvist, Flykt, & Öhman, 1998) from the Karolinska Institute, Stockholm, Sweden, (emotionally neutral photographs of individuals AM01, AM02, AM04, AM05, AM10, AM14, AM17, AM21, AM23, AM26, AM28, AM31 for men, and AF01, AF04, AF06, AF08, AF11, AF13, AF17, AF23, AF27, AF31, AF34, AF35 for women were selected).

From each selected photograph a (“neutral”) 3D morphable model was generated following the procedure suggested by Blanz and Vetter (1999) and Singular Inversions (2010) (see two examples in Figures 1a and 1b). All models were without head hair, makeup, or accessories.

From each "neutral" model, a "homosexual" and a "heterosexual" version were created following the same procedure as González-Álvarez (2017, Exp. 1).

Procedure.

Each participant responded to all pairs of faces (homosexual and heterosexual versions derived from the same "neutral" face) presented in random order in two sessions separated by a short break. The task was performed individually[1] online through the university intranet. In every trial, two faces were presented horizontally and the participant had to mark with the mouse the face that apparently corresponded to a person with more probabilities to have a homosexual orientation. Homosexual and heterosexual versions were presented randomly on the left or right side of the screen.

RESULTS AND DISCUSSION

Responses selecting the "homosexual" version were counted as correct. Participants achieved an accuracy level of 0.67, (SD = 0.17), 95% CI [0.63, 0.71], significantly above the chance level (0.50); t (75) = 8.91, p < .0001, although slightly lower than what was found by González-Álvarez (2017) for totally artificial faces (0.74). The accuracy mean for male faces was 0.73 (SD = 0.17), 95% CI [0.68, 0.76], t (75) = 11.65, p < .0001, and for female faces was 0.63 (SD = 0.26), [0.56, 0.62], t (75) = 4.24, p < .0001. Figure 2 displays the means separated by the sex of the participants. All means were significantly above the chance level (0.50), as shown by the bars indicating the corresponding standard error means (SEM). It is noteworthy that women showed a pronounced sensitivity to detecting the "homosexual" versions of the male faces, reaching a clearly higher level of accuracy than the other means.

A 2 (Face Gender) x 2 (Sex of Participants) ANOVA was carried out, and we made separate analyses across participants (F_1) and items (F_2). The analysis revealed that the Face Gender effect did not reach significance level through subjects, $F_1(1, 74) = 1.92$, $MS_e = 0.035$, $p = .260$, $\eta^2_p = .017$; but reached marginal significance through items, $F_2(1, 46) = 3.98$, $MS_e = 0.013$, $p = .052$, $\eta^2_p = .080$, indicating that male faces obtained a higher level of accuracy than female faces (0.72 vs. 0.62), although the eta partial squared (η^2_p) showed a medium effect size[2]. The Sex of Participants effect did not reach significance level through subjects (most likely due to the small number of male participants), $F_1(1, 74) = 2.28$, $MS_e = 0.058$, $p = .135$, $\eta^2_p = .030$; but clearly it did through items, $F_2(1, 46) = 24.16$, $MS_e = 0.006$, $p < .0001$, $\eta^2_p = .344$, suggesting that women (0.69) were better than men (0.61) at detecting “homosexual” faces.

The ANOVA revealed a Face Gender x Sex of Participants interaction both across subjects, $F_1(1, 74) = 5.17$, $MS_e = 0.035$, $p = .026$, $\eta^2_p = .065$, and across items, $F_2(1, 46) = 31.19$, $MS_e = 0.006$, $p < .0001$, $\eta^2_p = .404$, mainly caused because the high performance of women with the male faces.

Consequently, our data demonstrated that the ability to identify “homosexual” versions of face models created from photographs of real people was similar to what was found by González-Alvarez (2017) using artificial faces created *ex novo* at random by computer. In both cases, stimuli were manipulated by modifying the facial measures identified as predictors of sexual orientation (Skorska et al., 2015). The present results, along with those obtained by González-Alvarez (2017), add perceptual legitimacy and provide a cross-cultural consistency to the Skorska et al.’s findings.

Nevertheless, the present experiment was based on a binary task consisting of choosing which of two faces corresponded to a person with more probabilities to have a homosexual orientation. Data showed that participants were able to complete this task clearly above the chance level. The binary task used in this experiment involved an all-or-nothing decision that was not very complex from a perceptive/cognitive point of view, and it did not allow different gradations of the facial features studied. In the next experiment, we used a task based on a rating scale, that is more demanding from a perceptual/cognitive point of view.

EXPERIMENT 2

This experiment tested how people rate the apparent sexual orientation of 3D face models created from photographs of real people and manipulated according to the main multivariate variables reported by Skorska et al. (2015) as predictors of sexual orientation. The task used in this experiment was more demanding than mere binary discrimination between two face models. The main aim of the experiment was to test whether lay (young) people were sensitive to different gradations of the key facial attributes, and if their sexual orientation ratings were mostly categorical or continuous in nature. We hypothesized that participants will rate the apparent sexual orientation of realistic face models as a continuous linear function of the values of the manipulated variables.

METHODS

Participants.

In this experiment participated seventy-seven adults of both sexes (55 females); whose age range was 17-39 years (M = 20.03; SD = 3.85). All of them were undergraduates at the University xxxxxxxxxxxxxx, who participated voluntarily for course credit and none had participated in Experiment 1. Of those who indicated ethnicity (n=70), 81% were White/Caucasian, 16% were Hispanic/Latin American, and 3% were another ethnicity.

Materials.

The stimuli were 100 realistic three-dimensional face models created from photographs of real people in two steps. First, twenty photographs, ten of men and ten of women, were selected from the Basel Face Database (Walker et al., 2018; photographs of individuals 5, 9, 16, 25, 28, 29, 31, 32, 33, 40 for men, and 7, 8, 12, 13, 14, 23, 24, 30, 36, 39 for women; all Caucasian), and, through the FaceGen software, they were transformed into twenty 3D ("neutral") face models. Second, four additional faces (two "homosexual" versions and two "heterosexual" versions) were created from every "neutral" face by manipulating the same attributes manipulated in Experiment 1 and following the same procedure as González-Álvarez (2017, Exp. 2). Thus, from each original photography, we constructed a series of five stimuli, ranging on equal steps from a "heterosexual" pole or extreme (version 1) to a "homosexual" pole (version 7); see two examples in Figure 3. The male and female stimuli were presented in two separated sessions in random order. All faces were displayed at a yaw angle of 13º and a pitch angle of -8º.

Procedure.

Each participant responded to all the faces in random order and in two sessions, one for male faces and another for female faces, in counterbalanced order across the participants. The task was performed individually online through the university intranet.

In each trial, a single face was presented and the participant had to indicate on a scale how likely s/he believed that the person represented by the face had a homosexual orientation. The scale had seven marks ranging from "1-No or very little probability of homosexual orientation" to "7-Quite a high probability of homosexual orientation". There was no time restriction to respond.

RESULTS AND DISCUSSION

Figure 4 displays the rating means for the five versions of male and female faces. The values suggest that the perception of homosexuality from facial stimuli was continuous rather than categorial, and the rating scores almost perfectly fit a linear regression model of the five face versions, especially for male faces. The same pattern was also found by González-Álvarez (2017) for entirely artificial stimuli not based on faces of real people. In the case of the male faces, the correlation between ratings and the five versions was $r = 0.996$ ($p < .0001$), and the rating means perfectly fit ($R^2 = 0.9915$) a linear regression model represented by the equation $y = 0.279x + 2.811$. For the female faces, the correlation was $r = 0.963$ ($p = .009$), and the means reasonably fit ($R^2 = 0.9258$) a linear regression model represented by the equation $y = 0.3296x + 3.0937$. It should be noted that the fit of the female faces was less perfect than that of the male faces, and versions 1 and 2 of the female faces were not perceived as different in terms of their degree of "homosexuality".

As in González-Álvarez (2017), the female faces obtained a greater score of homosexuality than the male stimuli (Figure 4), with the constant part of the regression equation, or the y-intercept, being higher for female faces than for male faces (3.0937 vs. 2.811). The overall means also differed for both gender sets of stimuli; female faces: 4.08, 95% CI [3.95, 4.21]; male faces: 3.65, 95% CI [3.57, 3.77]. Moreover, as in González-Álvarez (2017), sensitivity to the five face versions was greater for the female stimuli than for the male stimuli: regression slope coefficients were 0.33 *vs.* 0.28, respectively.

Table 1 shows the rating means disaggregated by sex of participants and gender of faces. We carried out a 2 (Sex of Participant) x 2 (Face Gender) x 5 (Face Version) ANOVA and performed separate analyses through participants (F_1) and items (F_2). The analysis found a significant Face Gender effect both through subjects, $F_1(1, 75) = 19.33$, $MS_e = 1.487$, $p < .0001$, $\eta^2_p = .205$, and through items, $F_2(1, 90) = 11.07$, $MS_e = 0.832$, $p < .01$, $\eta^2_p = .110$, indicating that the female faces received higher rates than the male faces. However, the Sex of Participant factor did not reach significance level, suggesting that men and women did not differ as experimental raters: $F_1 < 1$; $F_2(1, 90) = 1.85$, $MS_e = 0.044$, $p = .177$, $\eta^2_p = .020$. As expected, Face Version yielded a robust effect both through subjects, $F_1(4, 300) = 101.86$, $MS_e = .556$, $p < .0001$, $\eta^2_p = .576$, and through items, $F_2(4, 90) = 11.79$, $MS_e = 0.832$, $p < .0001$, $\eta^2_p = .344$. Posterior within-subject contrasts revealed that the rate mean of every face version was significantly different (higher) from that the anterior version, except for versions 1 and 2 of female faces[3].

None of the possible interactions reached significance either through the subjects or through the items, except the Face Gender x Face Version interaction through subjects, $F_1(4, 75) = 3.94$, $MS_e = .456$, $p < .01$, $\eta^2_p = .050$, because the score differences between females vs. male faces do not remain the same value throughout the five versions (Table 4, Figure 4). Differences between female vs male faces were 0.52 for version 1; 0.26 for version 2; 0.24 for version 3; 0.49 for version 4; and 0.46 for version 5.

GENERAL DISCUSSION

An issue that has significant social implications and psychological interest is whether lay people can perceive sexual orientation at first sight. Gaydar, or the term that research literature typically defines as the ability to categorize or identify sexual orientation, is also a highly-sensitive public issue in today’s world (Miller, 2018). For example, the report of an artificial intelligence software that could identify gay men and lesbian women by their faces (Wang & Kosinski, 2018) opened a particularly intense debate in social media.

A valuable source of information in perceptive terms is the human face, even though few works have studied facial structure as a predictor of sexual orientation (Hughes & Bremme, 2011; Skorska et al., 2015; Valentova et al., 2014). The rationale behind this type of study is that some biological factors, such as prenatal or postnatal (puberty) exposure to certain levels of sex hormones, could subtly influence the configuration of the facial structure (differently in men and women) and, at the same time, their sexual orientation (for a recent review see Wang, Wu, & Sun, 2019).

The morphometric study by Skorska et al. (2015) with a large sample of Canadian people identified three facial anatomical parameters as multivariate predictors in gay/heterosexual men and four different parameters as predictors in lesbian/heterosexual women. An interesting further is to examine the perceptual validity of these anatomical predictors of sexual orientation. To this end, González-Álvarez (2017) created computer-generated artificial face models that varied only on these parameters, and he used them to assess their perception by Spanish individuals, that is, a sample of people belonging to a different cultural and geographical environment.

The stimuli used by González-Álvarez's (2017) were totally artificial faces created by computer. However, there is evidence of some processing differences between artificial and natural faces at the cognitive/emotional (Balas & Pacella, 2015; Carlson, et al., 2012; Dyck et al., 2008; Kätsyri & Sams, 2008; Philip, Martin, & Clavel, 2018; Salminen, Jung, Santos, & Jansen, 2020) and neural representation (Wheatley, et al., 2011) levels. In the present study, we replicated and extended González-Álvarez's work using face models created from photographs of real people. Our results confirm those found by González-Álvarez (2017), although the overall level of accuracy (0.67 in the discrimination experiment) has been somewhat lower than that obtained in that study (0.74). Both studies used the same procedure and kind of human sample

(Spanish university students), except for the administered materials. In the present work, the stimuli were created from photographs of real people, but devoid of any cultural cue (hair style, makeup, adornment etc.) that could be potentially informative about sexual orientation. These more ecologically valid stimuli are more complex and incorporate natural information on color and skin texture from the human models. This additional information *per se* is not informative about sexual orientation, but possibly adds "background noise" when making judgments about sexual orientation. On the contrary, purely artificial stimuli, created *ex novo* by computer, perhaps enhances the opportunity for the manipulated variables to exert more direct influence on perceptual judgments, while a part of the ecological validity is sacrificed.

On the other hand, women showed a higher level of accuracy (0.69) than men (0.61) in the binary hetero/homosexual judgment task. Principally due to the great success of women in discriminating between gay / straight faces of men (0.75). In González-Álvarez's (2017) first experiment, a slight superiority (5%) of women over men was also found in the perception of the sexual orientation of faces. This slightly differential pattern in favor of women is consistent with some sex differences found in face processing (Rehnman, 2007), particularly in the domain of social cognition (Proverbio, 2017). More specifically, Rule, Rosen, Slepian, and Ambady (2011) found that heterosexual women judged men's sexual orientation more accurately (than women's sexual orientation) when motivated to mate with a priming task based on romantic thoughts.

In the present study, sexual orientation of our participants (university students) was not asked for ethical reasons. We assume that our sample would include about 5-6% of LGB people, corresponding to the overall rate for Spain. The evidence suggests that gay and lesbian individuals often achieve higher rates of accuracy than heterosexual individuals, though this varies depending on specific cues (see the review of Rule & Alaei, 2016).). Rule and Alaei

(2016) concluded in their review for *Current Directions in Psychological Science* that "overall […] people showing a greater history of approach motivations toward gay and lesbian individuals (e.g., greater familiarity, less prejudice) tend to be the most accurate" (p. 446).

Data from the second experiment revealed that homosexuality perceived from the versions originated from natural photographs manipulating the Skorska et al's. (2015) multivariate predictors was continuous in nature and almost a perfect linear function of the values of such parameters. Basically, it was the same pattern obtained by González-Álvarez (2017) with purely artificial stimuli. Also, in the present study the female faces were seen as somewhat more homosexual than the male faces -i.e., they obtained a greater score of homosexuality than the male stimuli- (Figure 4). In addition, participants were slightly more sensitive to the manipulated variables in the female faces, just as the regression slope coefficients demonstrated: every step from one female version to the next one increased 0.33 points in perception of homosexuality (on a 1-7 scale), whereas this gradient was only 0.28 points in the case of male stimuli. This difference was likely due to the fact that the fourth predictor of homosexuality in women identified by the multivariate analysis of Skorska et al. (2015) was that lesbian women had marginally more masculine face shapes than heterosexual women. In the FaceGen software, used by Skorska et al. (2015) and us, that overall feature is measured and manipulated by means of a general gender control, which could exert a relatively stronger effect on the stimulus appearance. At the same time, this parameter is in line with a traditional -and secular- "gender inversion" stereotype of gay men presenting as more feminine than heterosexual men and lesbian women presenting as more masculine than heterosexual women (Rule, 2017).

However, it is worth noting that rating scores to the female stimuli adjusted something worse to a linear function than rating scores to the male stimuli ($R^2 = 0.9258$ vs. $R^2 = 0.9915$, respectively; see Figure 4). One of the reasons for this worst fit is that female versions 1 and 2 received roughly the same total punctuation of homosexuality (3.60). Examining the rating means disaggregated by sex of participants (Table 1), it is curious that the anomaly only occurs in the scores given by the female judges. As expected, men perceived the female face 1 as less homosexual (3.49) than the female face 2 (3.63). However, women perceived the female face version 1 - who had the most pronounced feminine features- as more homosexual (3.65) than the face version 2 (3.59). Nevertheless, any difference between male and female judgments should be interpreted with caution, given the small number of male participants.

As stated in the introduction, an additional reason for carrying out the present study had to do with the so-called reproducibility crisis in psychology (particularly in social psychology) because the Reproducibility Project reported in *Science* that only about half of the effects published in psychology journals are effectively replicated (Open Science Collaboration 2015). However, as Baucal et al. (2020) recently stated, we must not forget that many findings obtained in social and cultural psychology correspond to phenomena that can vary across different socio-cultural contexts. In any case, the present results robustly replicate González-Álvarez's (2017) findings, and they reaffirm, with new stimuli based on real faces, the perceptual validity of the seven multivariate facial parameter predictors of sexual orientation identified by Skorska et al (2015).

In summary, the two experiments included in this study have confirmed the robustness of the previous findings and extended the investigation by experimentally manipulating the key facial parameters in 3D models created from photographs of real people. Data from Experiment 1 showed an overall accuracy (0.67) significantly above the chance level in a binary hetero/homosexual judgment task. Data from the Experiment 2 revealed that participants rated the apparent sexual orientation of series of 3D face models as a continuous linear function of the manipulated key parameters. Thus, this study contributed to demonstrate the perceptual validity of the seven multivariate predictors identified by Skorska et al. (2015).

As mentioned above, a limitation of the present study is the relatively small number of male participants in relation to that of women. It is the usual consequence of a very pronounced gender asymmetry always existent in the university colleges of Psychology in Spain and other countries. In the future, it would be interesting to deepen into the gender differences that emerge in the perception of sexual orientation with more balanced samples between men and women. Another issue of further research would be to explore the perception of sexual orientation if all 11 (for men) and 17 (for women) Skorska et al's (2015) univariate facial metrics were manipulated. Furthermore, from a methodological point of view, the present study, together with the one by González-Álvarez (2017), opens up new avenues for further research on the facial structure/sexual orientation interaction by experimentally manipulating face models.

Finally, it should be noted that this study, as many other gender/sex studies, uses binary conceptualizations of sexual orientation (as well as gender/sex), opposing "heterosexual" against "homosexual" and obscuring other sexual orientations with regard to both facial features and recognition via facial structures. This issue could be an important avenue for future research.

Footnotes

[1] We assume that the administration was individual. The instructions clearly indicated that the experiment was to be performed individually in classroom or a quiet environment without distraction. This could be a limitation of the procedure.

[2] The effect size interpretations for η^2_p values are: .01 = small, .06 = medium, and .14 = large.

[3]. Rating differences between consecutive face versions yielded the following p-values. For male faces: versions 1 vs. 2 ($p = .00009$); 2 vs. 3 ($p <.00001$); 3 vs. 4 ($p <.00001$); 4 vs. 5 ($p = .00058$). For female faces: versions 1 vs. 2 ($p = .97820$); 2 vs. 3 ($p <.00001$); 3 vs. 4 ($p <.00001$); 4 vs. 5 ($p = .00001$);

Ethical Compliance Section

This work was completed with resources provided by the University Jaume I (Spain).

All procedures performed in studies involving human participants were in accordance with the Deontological commission and of the Ethical Committee of Animal Welfare of the University Jaume I (Spain) and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

The authors declare they have no conflict of interest.

Written informed consent was obtained from all individual adult participants included in the study.

References

Balas, B., & Pacella, J. (2015). Artificial faces are harder to remember. *Computers in human behavior*, *52*, 331-337. doi:10.1016/j.chb.2015.06.018.

Baucal, A., Gillespie, A., Krstić, K., & Zittoun, T. (2020). Reproducibility in psychology: Theoretical distinction of different types of replications. *Integrative Psychological and Behavioral Science*, *54*(1), 152-157. doi:10.1007/s12124-019-09499-y.

Bjornsdottir, R. T., & Rule, N. O. (2020). Emotion and Gender Typicality Cue Sexual Orientation Differently in Women and Men. *Archives of Sexual Behavior*, *49*, 2547-2560. doi.org/10.1007/s10508-020-01700-3

Blanz, V., & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. *In Proceedings of the 26th annual conference on Computer graphics and interactive techniques* (pp. 187-194). New York, July, 1999. doi:10.1145/311535.311556.

Carlson, C. A., Gronlund, S. D., Weatherford, D. R., & Carlson, M. A. (2012). Processing differences between feature-based facial composites and photos of real faces. *Applied Cognitive Psychology*, *26*(4), 525-540. doi:10.1002/acp.2824.

Dyck, M., Winbeck, M., Leiberg, S., Chen, Y., Gur, R. C., & Mathiak, K. (2008). Recognition profile of emotions in natural and virtual faces. *PLoS One*, *3*(11), e3628. doi:10.1371/journal.pone.0003628.

Ehrlich, S. M., Schiano, D. J., & Sheridan, K. (2000, April). Communicating facial affect: it's not the realism, it's the motion. In *CHI'00 extended abstracts on human factors in computing systems* (pp. 251-252). doi:10.1145/633292.633439.

Freeman, J. B., Johnson, K. L., Ambady, N., & Rule, N. O. (2010). Sexual orientation perception involves gendered facial cues. *Personality and Social Psychology Bulletin*, *36*, 1318-1331. doi:10.1177/0146167210378755.

González-Alvarez, J. (2017). Perception of sexual orientation from facial structure: A study with artificial face models. *Archives of Sexual Behavior*, *46*(5), 1251-1260. doi:10.1007/s10508-016-0929-6.

González-Álvarez, J., & Cervera-Crespo, T. (2019). Gender differences in sexual attraction and moral judgment: research with artificial face models. *Psychological Reports*, *122*(2), 525-535. doi:10.1177/0033294118756891

Hughes, S. M., & Bremme, R. (2011). The effects of facial symmetry and sexually-dimorphic facial proportions on assessments of sexual orientation. *Journal of Social, Evolutionary, and Cultural Psychology*, *5,* 214–230. doi:10.1037/h0099261.

Kätsyri, J., & Sams, M. (2008). The effect of dynamics on identifying basic emotions from synthetic and natural faces. *International Journal of Human-Computer Studies*, *66*(4), 233-242. doi:10.1016/j.ijhcs.2007.10.001.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.

Miller, A. E. (2018). Searching for gaydar: Blind spots in the study of sexual orientation perception. *Psychology & Sexuality, 9*(3), 188-203. doi:10.1080/19419899.2018.1468353.

Open Science Collaboration (2015). Psychology. Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. doi:10.1126/science.aac4716

Philip, L., Martin, J. C., & Clavel, C. (2018). Rapid facial reactions in response to facial expressions of emotion displayed by real versus virtual faces. *i-Perception*, *9*(4), 2041669518786527. doi:10.1177/2041669518786527

Proverbio, A. M. (2017). Sex differences in social cognition: The case of face processing. *Journal of neuroscience research*, *95*(1-2), 222-234. doi:10.1002/jnr.23817

Rehnman, J. (2007). *The role of gender in face recognition* (Doctoral dissertation, Stockholm University).

Rule, N. O. (2016). Perceptions of sexual orientation from minimal cues. *Archives of Sexual Behavior*. Advance online publication. doi:10.1007/s10508-016-0779-2

Rule, N. O., & Alaei, R. (2016). "Gaydar" the perception of sexual orientation from subtle cues. *Current Directions in Psychological Science*, *25*(6), 444-448.

Rule, N. O., & Ambady, N. (2008). Brief exposures: Male sexual orientation is accurately perceived at 50 ms. *Journal of Experimental Social Psychology*, *44*, 1100-1105. doi:10.1016/j.jesp.2007.12.001.

Rule, N. O., Ambady, N., Adams, R. B., & Macrae, C. N. (2008). Accuracy and awareness in the perception and categorization of male sexual orientation. *Journal of Personality and Social Psychology*, *95*, 1019-1028. doi:10.1037/a0013194.

Rule, N. O., Macrae, C. N., & Ambady, N. (2009). Ambiguous group membership is extracted automatically from faces. *Psychological Science*, *20*, 441-443. doi:10.1111/j.1467-9280.2009.02314.x.

Rule, N. O., Rosen, K. S., Slepian, M. L., & Ambady, N. (2011). Mating interest improves women's accuracy in judging male sexual orientation. *Psychological Science*, *22*(7), 881-886.

Salminen, J., Jung, S. G., Santos, J. M., & Jansen, B. J. (2020). Does a Smile Matter if the Person Is Not Real?: The Effect of a Smile and Stock Photos on Persona Perceptions. *International Journal of Human–Computer Interaction*, *36*(6), 568-590. doi:10.1080/10447318.2019.1664068.

Skorska, M. N., Geniole, S. N., Vrysen, B. M., McCormick, C. M., & Bogaert, A. F. (2015). Facial structure predicts sexual orientation in both men and women. *Archives of Sexual Behavior*, *44*, 1377-1394. doi:10.1007/s10508-014-0454-4.

Tskhay, K. O., & Rule, N. O. (2013). Accuracy in categorizing perceptually ambiguous groups: A review and meta-analysis. *Personality and Social Psychology Review*, *17*, 72-86. doi:10.1177/1088868312461308.

Valentova, J. V., Kleisner, K., Havlicek, J., & Neustupa, J. (2014). Shape differences between the faces of homosexual and heterosexual men. *Archives of Sexual Behavior*, *43*, 353–361. doi:10.1007/s10508-013-0194-x.

Valentova, J. V., Rieger, G., Havlicek, J., Linsenmeier, J. A. W., & Bailey, J. M. (2011). Judgments of sexual orientation and masculinity-femininity based on thin slices of behavior: A cross-cultural comparison. *Archives of Sexual Behavior*, *40*, 1145-1152. doi:10.1007/s10508-011-9818-1.

Walker, M., Schönborn, S., Greifeneder, R., & Vetter, T. (2018). The Basel Face Database: A validated set of photographs reflecting systematic differences in Big Two and Big Five personality dimensions. PloS one, 13(3), e0193190. doi:10.1371/journal.pone.0193190

Wang, Y., & Kosinski, M. (2018). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of personality and social psychology*, *114*(2), 246. doi:10.1037/pspa0000098.

Wang, Y., Wu, H., & Sun, Z. S. (2019). The biological basis of sexual orientation: How hormonal, genetic, and environmental factors influence to whom we are sexually attracted. *Frontiers in Neuroendocrinology*, *55*, 100798. doi:10.1016/j.yfrne.2019.100798

Wheatley, T., Weinberg, A., Looser, C., Moran, T., & Hajcak, G. (2011). Mind perception: Real but not artificial faces sustain neural activity beyond the N170/VPP. *PloS one*, *6*(3), e17960. doi:10.1371/journal.pone.0017960

Figure 1a. Examples of stimuli (C and D) presented in Experiment 1. A: Photograph of a man selected from the Basel Face Database (Walker, Schönborn, Greifeneder, & Vetter, 2018). B: Three-dimensional “neutral” model (not presented in the experiment) created from the photograph by means of the FaceGen Modeller software. C: “Heterosexual” version created from the “neutral” model. D: “Homosexual” version created from the “neutral” model.

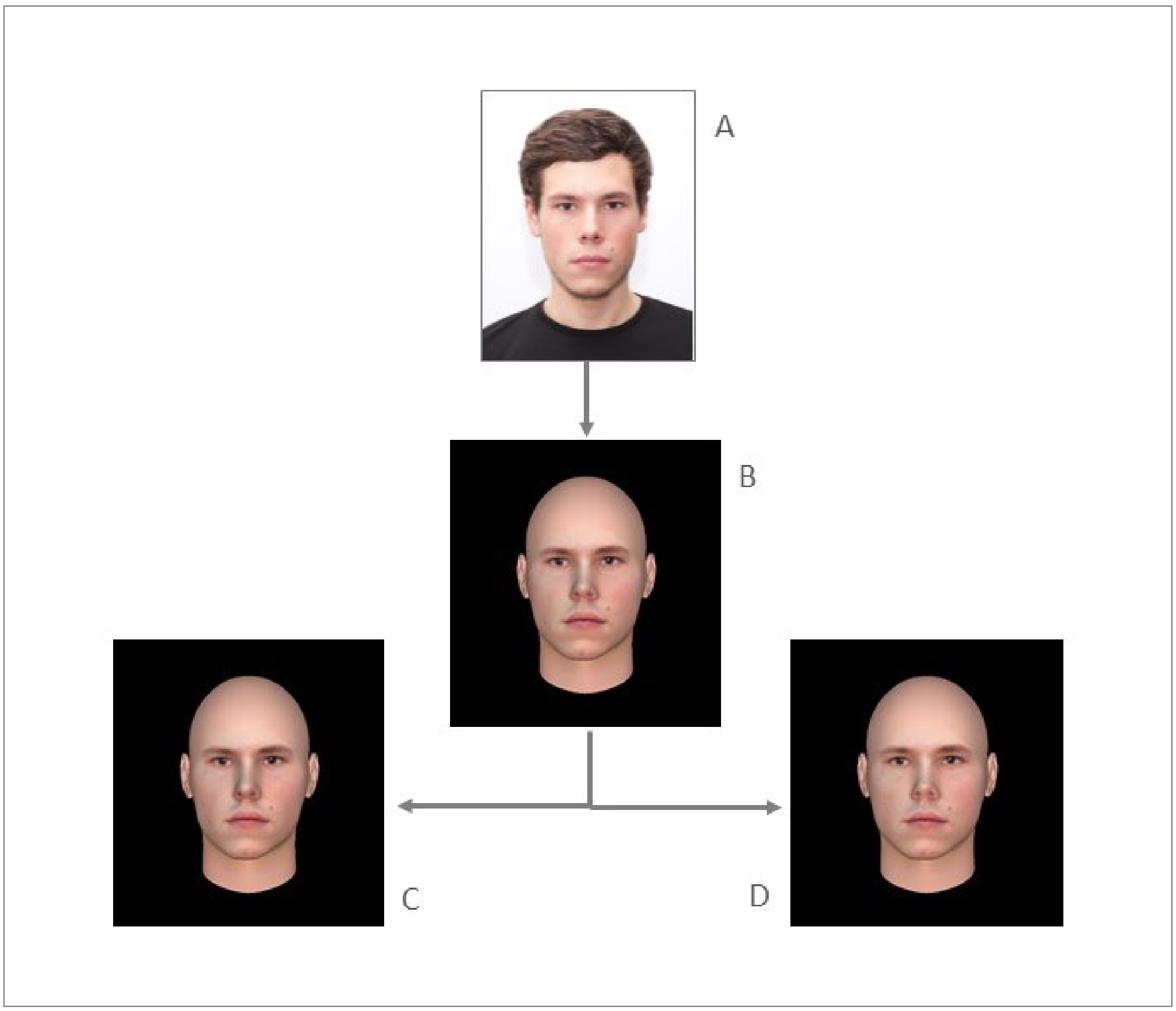

Figure 1b. Idem Figure 1a. Examples of stimuli from a photograph of a woman selected from the same face database.

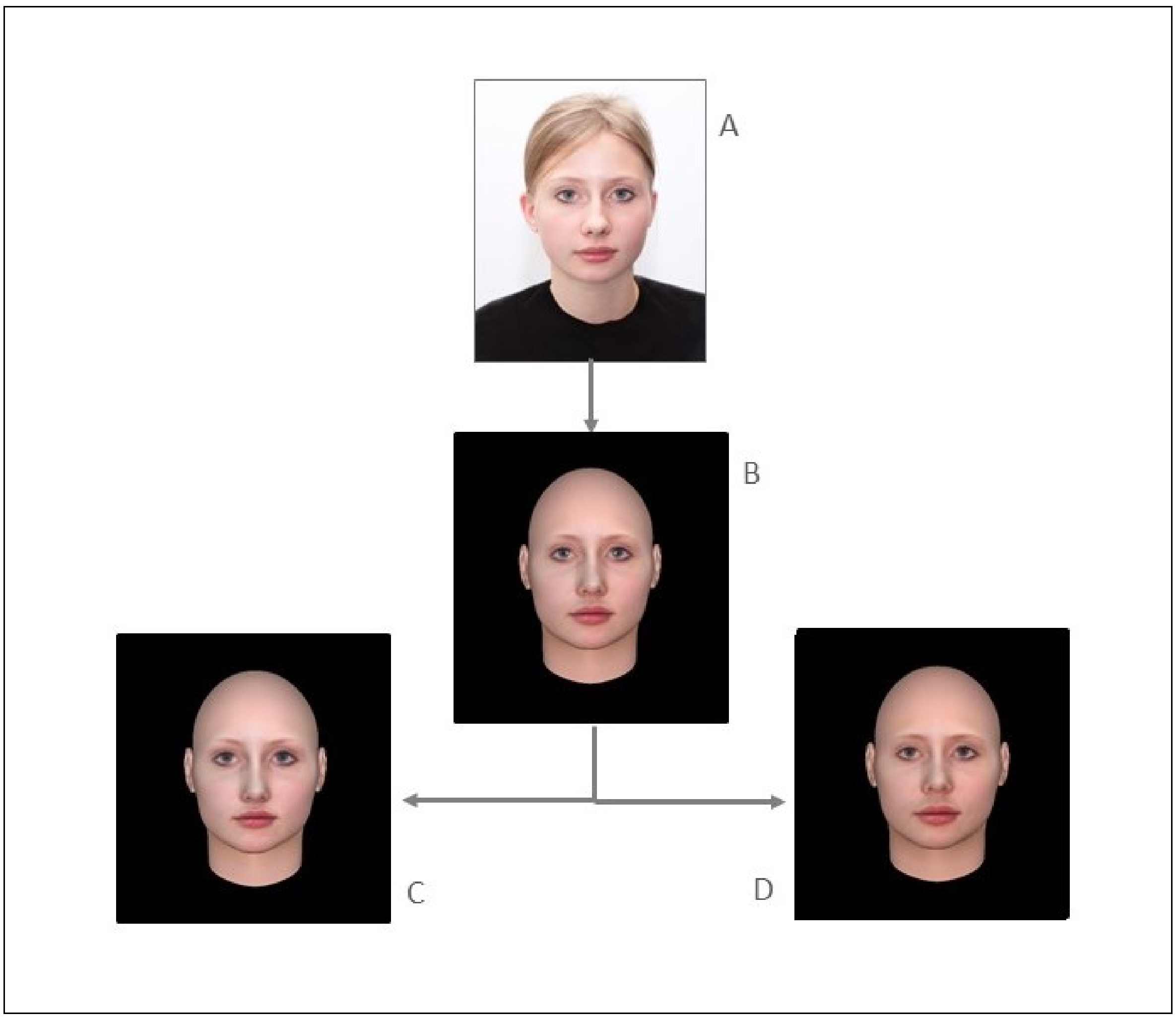

Figure 2. Data from Experiment 1. Proportion of correct responses to male and female faces by male and female raters. Error bars specify ± SEM (standard error of the mean).

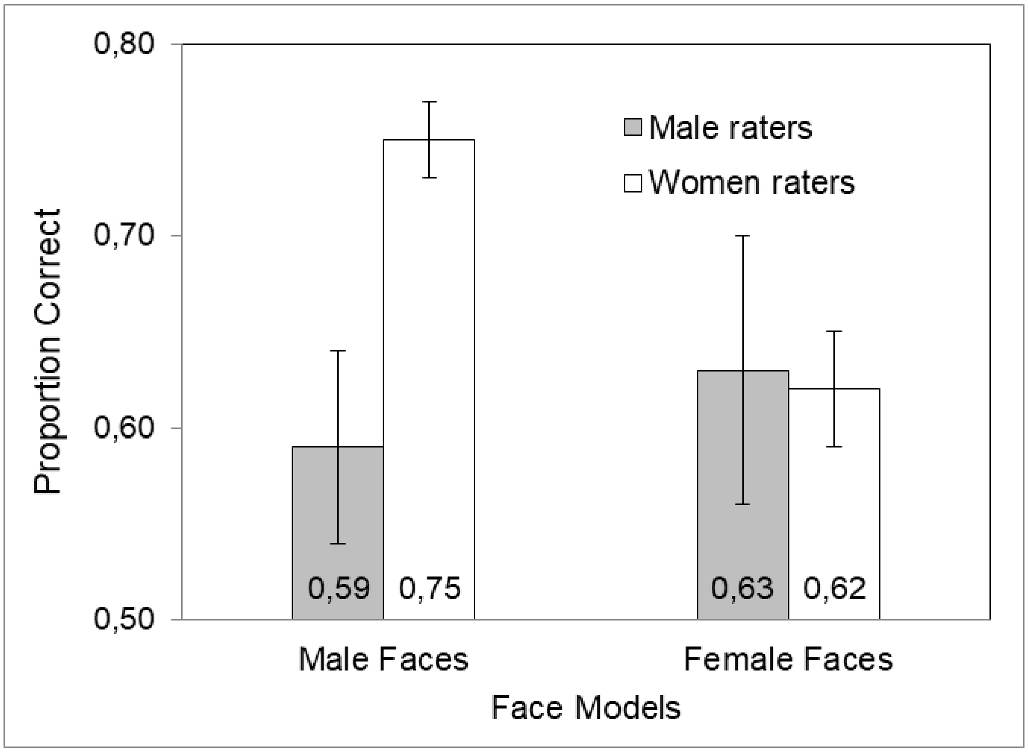

Figure 3. Examples of stimuli presented in Experiment 2, created from a photograph of two real persons (man: upper panel; woman: lower panel). Faces range in equal steps from a "heterosexual" pole (1) to a "homosexual" (5) pole. More details in the text.

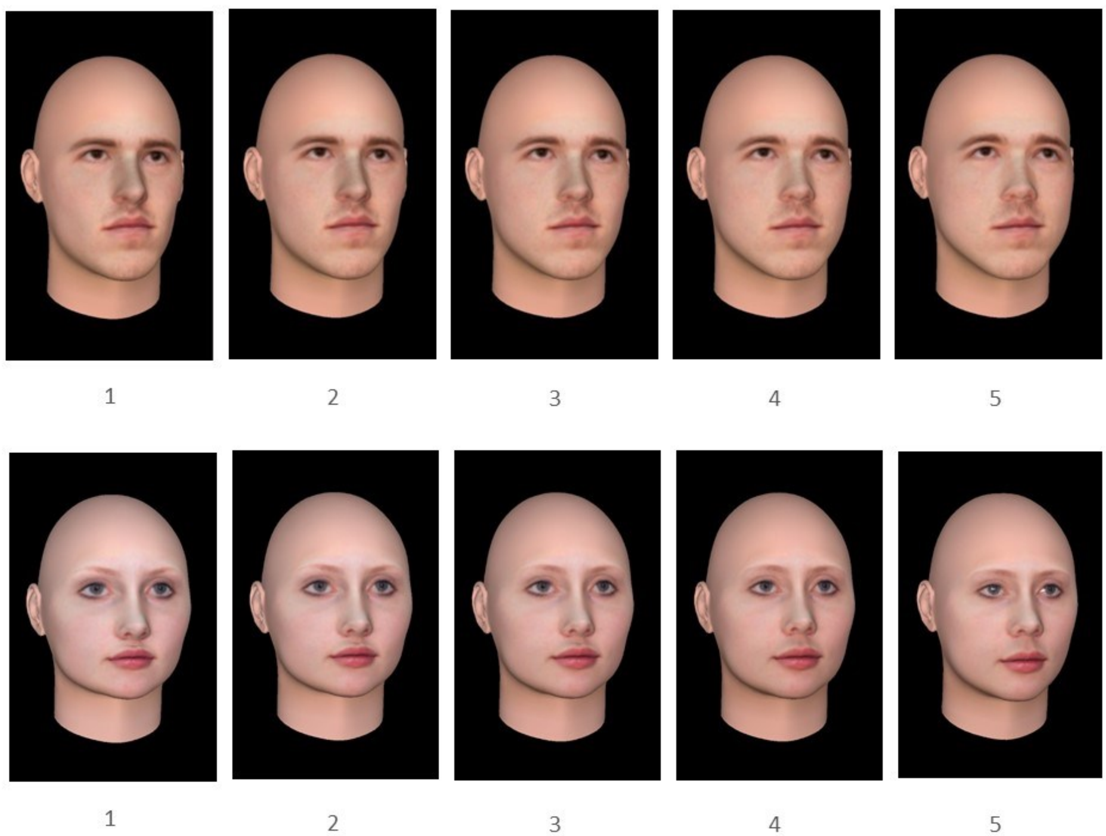

Figure 4. Experiment 2. Equations and regression lines between the rating scores and the five face versions. Data from male and female faces are ~~s~~eparated.

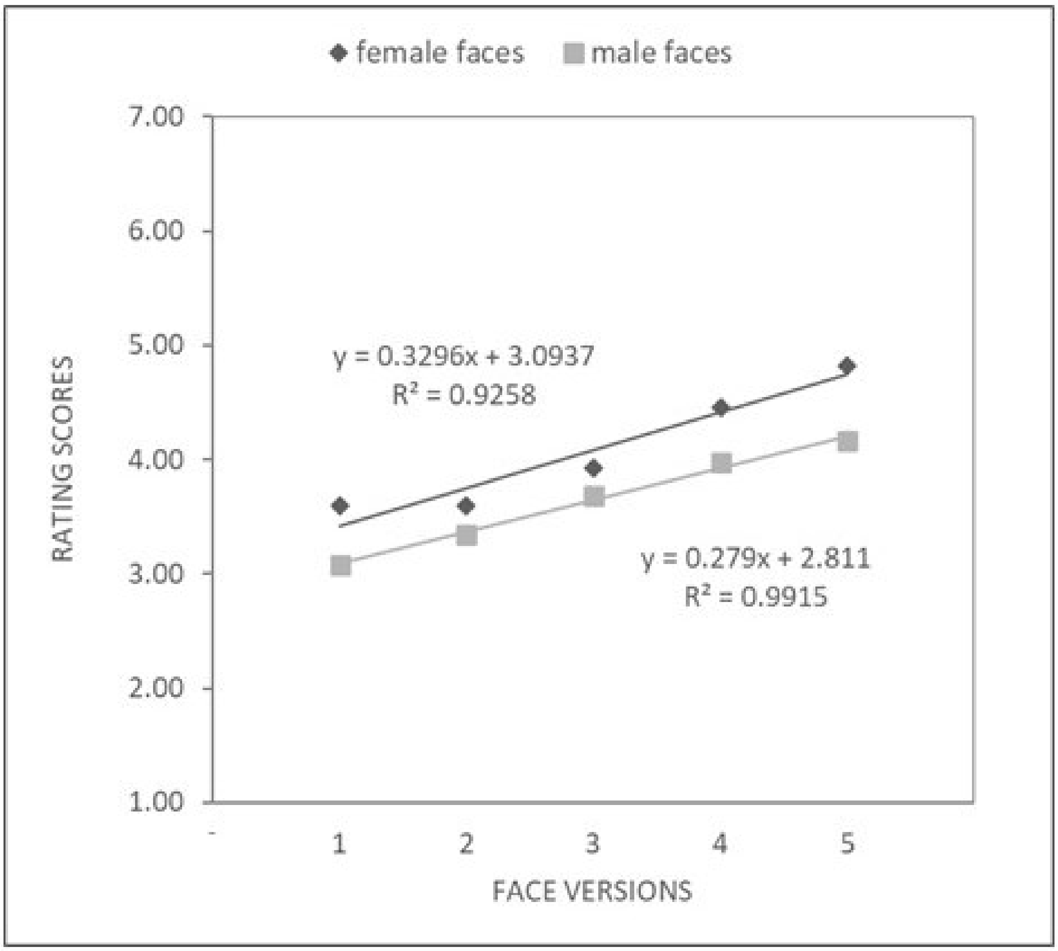

Table 1. Data from Experiment 2. Rating means (SD in parentheses) for the five versions of male and female faces disaggregated by the sex of the participants.

	Male faces (versions)					Female faces (versions)				
Sex of participants:	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Men:	3.22 (0.56)	3.35 (0.63)	3.69 (0.60)	3.97 (0.68)	4.23 (0.73)	3.49 (0.60)	3.63 (0.68)	3.87 (0.73)	4.60 (0.75)	4.89 (0.99)
Women:	3.02 (0.54)	3.34 (0.58)	3.69 (0.58)	3.97 (0.76)	4.13 (0.74)	3.65 (0.78)	3.59 (0.70)	3.95 (0.71)	4.41 (0.89)	4.79 (1.09)
Total:	3.08 (0.55)	3.34 (0.59)	3.69 (0.58)	3.97 (0.74)	4.16 (0.73)	3.60 (0.73)	3.60 (0.69)	3.93 (0.71)	4.46 (0.85)	4.82 (1.06)