Arch Sex Behav
DOI 10.1007/s10508-016-0929-6

ORIGINAL PAPER

Perception of Sexual Orientation from Facial Structure: A Study with Artificial Face Models

Julio González-Álvarez[1]

Received: 19 July 2015 / Revised: 21 September 2016 / Accepted: 26 December 2016

Abstract Research has shown that lay people can perceive sexual orientation better than chance from face stimuli. However, the relation between facial structure and sexual orientation has been scarcely examined. Recently, an extensive morphometric study on a large sample of Canadian people (Skorska, Geniole, Vrysen, McCormick, & Bogaert, 2015) identified three (in men) and four (in women) facial features as unique multivariate predictors of sexual orientation in each sex group. The present study tested the perceptual validity of these facial traits with two experiments based on realistic artificial 3D face models created by manipulating the key parameters and presented to Spanish participants. Experiment 1 included 200 White and Black face models of both sexes. The results showed an overall accuracy (0.74) clearly above chance in a binary hetero/homosexual judgment task and significant differences depending on the race and sex of the face models. Experiment 2 produced five versions of 24 artificial faces of both sexes varying the key parameters in equal steps, and participants had to rate on a 1–7 scale how likely they thought that the depicted person had a homosexual sexual orientation. Rating scores displayed an almost perfect linear regression as a function of the parameter steps. In summary, both experiments demonstrated the perceptual validity of the seven multivariate predictors identified by Skorska et al. and open up new avenues for further research on this issue with artificial face models.

Keywords Sexual orientation · Facial structure · Artificial faces · Homosexuality · Perception

✉ Julio González-Álvarez
gonzalez@psb.uji.es

[1] Department of Basic and Clinical Psychology and Psychobiology, University of Jaume I, Avinguda Sos Baynat s/n, 12071 Castellón de la Plana, Castellón, Spain

Introduction

An important issue in person perception is the categorization of people into perceptually ambiguous groups in the absence of obvious clues. Age, sex, race, for example, are easily inferred when two individuals meet for the first time, but other features such as professions, religious/political affiliations, or sexual orientation are perceptually elusive. The term "gaydar" (a linguistic blend of *gay* and *radar*) refers to the popular belief that heterosexual and gay/lesbian persons can be intuitively distinguished on the basis of different and mainly nonverbal cues (Rule, 2017). Perceptual sensitivity to sexual orientation may play an important role in human sexual behavior and is likely part of a refined mate-recognition system that helps to find potential mating partners. Actually, this issue is also relevant for its numerous social implications regarding the respect and civil rights of gays and lesbians. For example, it is good news that the Obama administration finally repealed in 2010 the "don't ask, don't tell" policy of the U.S. Army instituted in 1994, but many homosexual persons still experience (sometimes subtle) discrimination in everyday life on the basis of their sexual orientation and, what is worse, most of them suffer homophobia and persecution in many countries around the world.

Recently, Tskhay and Rule (2013) carried out a review and meta-analysis on the accuracy in categorizing perceptually ambiguous groups—Jewish group membership, religious affiliation (Mormon), political orientation (Democrat vs. Republicans), and sexual orientation—and they found an overall moderate-to-small significant correlation of $r = .29$ between perceived and actual (self-reported) sexual orientation. The review revealed that experimental participants can identify sexual orientation (heterosexual vs. gay/lesbian) better than chance from video stimuli of different lengths (Ambady, Hallahan, & Conner, 1999; Berger, Hank, Rauzi, & Simkins, 1987; Valentova, Rieger, Havlicek, Linsenmeier, & Bailey, 2011), still images, and even

Published online: 02 February 2017

from very brief exposures (50 ms) of photographs (Rule & Ambady, 2008). It seems that people apply stereotypes of gender inversion (Freeman, Johnson, Ambady, & Rule, 2010) and use sex atypicality as a cue to identify the homosexual orientation (Rieger, Linsenmeier, Gygax, García, & Bailey, 2010), but research indicates that these sources are not the only cues used by perceivers (Freeman et al., 2010). Interestingly, homosexual participants could be identified better than chance even though they were instructed to conceal their sex-atypical behaviors (Sylva, Rieger, Linsenmeier, & Bailey, 2010).

Human faces are an important source of personal information, and evidence suggests that sexual orientation can be inferred better than chance from a whole face (e.g., Freeman et al., 2010; Rule & Ambady, 2008) or from separate facial features (Rule, Ambady, Adams, & Macrae, 2008; Rule, Ambady, & Hallett, 2009a). Furthermore, priming-based data show that such inferences occur automatically (Rule, Macrae, & Ambady, 2009b).

It is true that facial-based judgments may exploit obvious cues such as hairstyle and other features that explicitly communicate nonverbal information about the person's sexual orientation, but experimental results suggest that lay people can make accurate and intuitive judgments on the basis of non-obvious and subtle information associated to certain anatomical elements of face (eyes, mouth area, etc.) (Rule et al., 2008). In this sense, a previous issue would be whether sexual orientation is actually associated to certain features of facial structure. So far, only three studies have investigated the facial structure in relation to sexual orientation.

Hughes and Bremme (2011) examined 60 photographs obtained from public open-access social networking profile pages, where individuals of both sexes had stated their sexual orientation. They found that self-identified heterosexuals had more symmetrical facial measures than homosexuals, but both groups did not differ in a set of seven sexually dimorphic facial measures (eye size, lower face/face height, cheekbone prominence, face width/lower face height, mean eyebrow height, forehead height, and lip/jaw width). Nevertheless, an examination of a composite score of these seven traits showed that heterosexual men had greater overall masculine facial features than gay men but no association was found between the composite measure and sexual orientation in women.

A second study by Valentova, Kleisner, Havlicek, and Neustupa (2014) in the Czech Republic tested the possible differences in facial shape between 40 heterosexual and 40 homosexual men. A morphometric analysis based on facial photographs taken from the participants revealed that gay men had relatively wider and shorter faces, more rounded jaws, and smaller and shorter noses, which resulted in a mixture of both feminine and masculine features.

Lastly, Skorska, Geniole, Vrysen, McCormick, and Bogaert (2015) carried out an extensive study in Canada with 129 homosexual and 261 heterosexual persons of both sexes. A facial photograph was taken from each individual, which served as input in a facial modeling program (Singular Inversions, 2010). After inputting a photograph into FaceGen, the program provides 62 facial metrics using statistical algorithms developed from 3D laser scans of human faces. Sixty-one of these parameters have numerical values expressed in standardized units, grouped into 10 featural categories (brow, cheek, mouth, nose, jaw, etc.). First, the data were analyzed at the univariate level to examine the partial correlations between each facial parameter and sex (men vs. women) and between each facial parameter and sexual orientation (homosexual vs. heterosexual). At this univariate level, lesbian and heterosexual women differed in 17 facial parameters or traits, while gay and heterosexual men differed in 11 facial traits. Note that some, but not all, of these parameters differed between the sexes. In a second step, Skorska et al. submitted the data to multivariate analysis and identified three unique multivariate predictors of sexual orientation within males and four within females. Concretely, homosexual men had more convex cheeks, shorter noses (as in heterosexual women), and foreheads that tilted back more. Lesbian women had more turned up noses (as in heterosexual men), mouths that were more puckered (less retracted), smaller foreheads, and marginally more masculine face shapes (also in heterosexual men).

A question that emerges from Skorska et al.'s (2015) work is whether the main anatomical traits statistically associated to sexual orientation could influence the perception of a person's sexual orientation by lay people. This question could be answered by means of the experimental manipulation of such traits using artificial faces as stimuli. Considering that those features were identified from a large sample of Canadian people, a second question concerns to what extent their hypothetical perceptual relevance may be generalizable beyond the geographical and cultural environment. At the same time, keeping in mind that only White subjects were included within the Canadian sample, one wonders if those facial features could preserve their informative value in the context of another race, concretely being part of a Black (African) face. This last issue has a double interest because from a perceptual point of view, it would test if those features exhibit certain cross-race generalizability, and, unfortunately, the identification of homosexual sexual orientation of an individual could result in negative consequences in many countries of Africa. For example, the recent "Uganda Anti-Homosexuality Act" (2014), known as the "Kill the Gays bill," initially proposed the death penalty for homosexuals in the original version and, currently, homosexuality is outlawed in 34 African nations (Ferreira, 2015).

The objective of the present study was to test the perceptive validity of the parameters identified by Skorska et al. (2015) as main predictors of sexual orientation, by means of two experiments using artificial faces as stimuli. For this purpose, the study collected three characteristics: creating 3D facial models, which varied only in the relevant metrics (3 in men, 4 in women) identified in the multivariate analysis of Skorska et al.; examining the race effect in the first experiment, including White (Caucasian) and Black (African) face models; and perceptually testing stimuli with (Spanish) participants who belonged to a different geographical and cultural background.

Experiment 1

The aim of the first experiment was to test whether people were able to discriminate between two artificial faces which differed only in the relevant metrics identified by Skorska et al. (2015) as predictors of sexual orientation in both men and women, and also in both Black and White faces. The research conformed to the American Psychological Association's Ethical Principles of Psychologists and Code of Conduct.

Method

Participants

Forty-nine young adults of both sexes participated in this experiment (35 females), whose age range was 18–35 years ($M = 21.00$; $SD = 2.81$). They were all undergraduates at the University Jaume I of Castellón (Spain), who volunteered in exchange for course credit. Of those who indicated ethnicity ($n = 40$), 99% were White/Caucasian and 1% was Hispanic/Latin American.

Materials

The experimental stimuli consisted of 200 realistic 3D artificial face models generated with the FaceGen Modeller 3.5 software (Singular Inversions, 2010): 50 of White males, 50 of White females, 50 of Black males, and 50 of Black females. All the facial models displayed the same default 3D position: yaw angle: 20.05° and pitch angle: 0.00° (see examples in Fig. 1).

White male faces were created as follows. First, FaceGen generated randomly 25 "neutral" faces of European males (these neutral faces were not used as stimuli). In a second step, from every neutral face, a "gay" and a "heterosexual" version were created by manipulating the following parameters that corresponded to the three predictors (gay men vs. heterosexual men) identified in the multivariate analysis of Skorska et al. (2015): Cheeks—concave/convex, Nose—short/long, and Forehead—tilt forward/back. The parameters were manipulated in ±2 standardized units (within a total range of 20 units).[1] FaceGen created Black male faces in the same way as White male faces, except that in the first step, the race control was set to the African racial origin position.

White female faces were created as follows. First, FaceGen generated randomly 25 "neutral" faces (not included as experimental stimuli) of European females. In a second step, from every neutral face a "lesbian" and a "heterosexual" version were created by manipulating the following parameters that corresponded to the four predictors (lesbian women vs. heterosexual women) identified in the multivariate analysis of Skorska et al. (2015): Nose—down/up, Mouth—protruding/retracted, Forehead—small/large, and General gender control. The first two parameters were manipulated in ±2 standardized units and the third parameter in ±1 unit[2] and the fourth parameter in ±8 steps (general gender control is different from the other controls and has a total of 80 steps). FaceGen created Black female faces in the same way as White female faces, except that in the first step, the race control was set to the African racial origin position.

Procedure

Each participant saw all the generated faces in random order and in four separated counterbalanced blocks in a within-subjects design. The task was completed individually online through the university intranet (virtual classroom). Previous research on face perception has demonstrated that laboratory and online studies produce equivalent results (e.g., DeBruine, Jones, Unger, Little, & Feinberg, 2007; Lefevre, Ewbank, Calder, von dem Hagen, & Perrett, 2013).

Participants wrote their name and demographic data and received the following instructions (in Spanish): "Recent research suggests that there may be subtle differences in facial structure associated with sexual orientation in both men and women. This is a perceptual study which extends that research. During each trial, two artificial faces created by computer will be displayed. Your task will consist in indicating which of the two faces you believe corresponds to a person who most likely has a homosexual orientation (i.e., toward the same sex). Here we use the term homosexual in a broad sense for both women and men."

During each trial, the computer displayed two faces horizontally, side by side, labeled A (left side) and B (right side): one was a "homosexual" version and the other was a "heterosexual" version, and both derived from the same "neutral" face. A question appeared below the faces: "Which one of these two faces do you think corresponds to a person who is more likely to have a homosexual orientation (A or B)?" Participants had to choose between two options (A or B) placed vertically by ticking one of them with a mouse. The side of presentation (left vs. right) was balanced across the trials between both face versions.

Each participant completed two sessions, separated by at least 1 h. The first session consisted of 25 trials that corresponded to a race-gender group (e.g., Black women), followed by 25 trials of the other gender from the same race group (Black men). The second session comprised 25 trials of one gender from the other race group (White men), followed by 25 trials of the other gender (White women). An open-ended question appeared at the end of every race-gender set of trials: "What did you base your answers on?"

[1] A previous pilot study showed that manipulation of ±2 standardized units in those parameters was sufficient to originate some (subtle) facial changes in a realistic way.

[2] A previous pilot study had shown that adding (or subtracting) two standardized units at the zero position of the Forehead—small/large control caused a more pronounced effect than in the other controls.

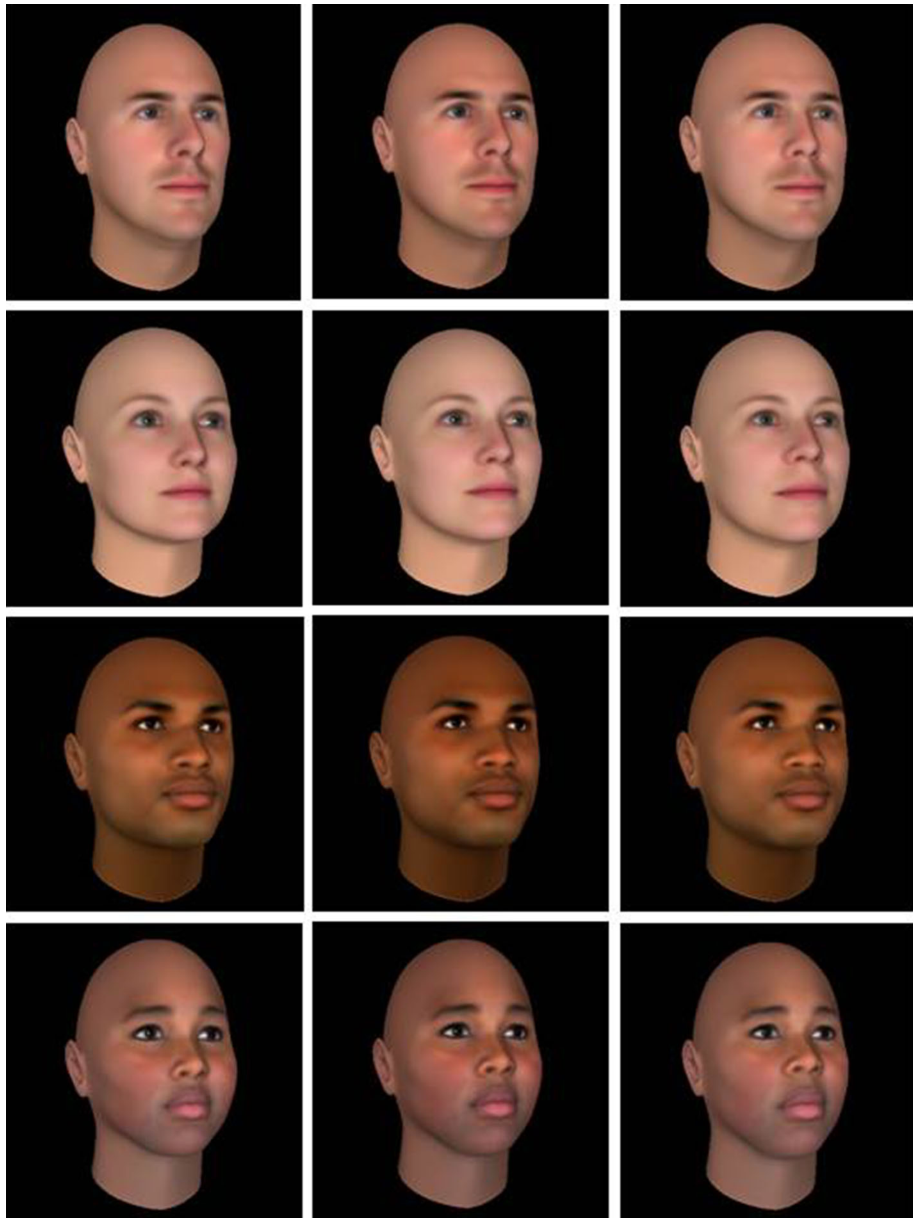

Fig. 1 Examples of the artificial faces created in Experiment 1. From the first to the fourth row: White men, White women, Black men, Black women, respectively. In each row, the central stimulus (not included in the experimental set) is a "neutral" face randomly generated with the FaceGen software; the left face is a "heterosexual" version that derived from the neutral face, and the right face is a "homosexual" version that derived from the neutral face

Results and Discussion

The responses that chose the "homosexual" face version were scored as correct. For several unforeseen reasons, seven participants (all females) did not complete the White faces set, and one participant (male) did not complete the Black faces set.

On average, participants reached an accuracy level of 0.74, 95% (SD = 0.16), CI [0.69, 0.79], which is clearly above the chance level (0.50). The accuracy means for each gender-race set of stimuli were the following: White male faces: 0.81 (SD = 0.15), 95% CI [0.76, 0.86]; White female faces: 0.76 (SD = 0.18), [0.70, 0.82]; Black male faces: 0.69 (SD = 0.17), [0.64, 0.74]; and Black female faces: 0.74 (SD = 0.25), [0.67, 0.81]. Figure 2 shows the means according to the sex of the participants. Interestingly, most participants did not identify the specific manipulated traits. The responses to the open-ended question focused on overall facial appearance; participants recognized "gay" faces because they looked "more feminine," "softer," "with fewer sharp features," "more peaceful," etc., or according to less precise statements, such as "I know intuitively," "based on first impressions," etc.; participants mainly recognized "lesbian" faces because they looked "more masculine," "had harder features," etc., or imprecisely, "it was my first impression," "they look like some homosexual women I know," etc. Very few people indicated some of the manipulated specific traits in men or women faces ("shorter nose," "chubby cheeks," etc.).

A 2 (Sex of Rater) × 2 (Face Gender) × 2 (Face Race) mixed-model analysis of variance (ANOVA) was conducted. Separate

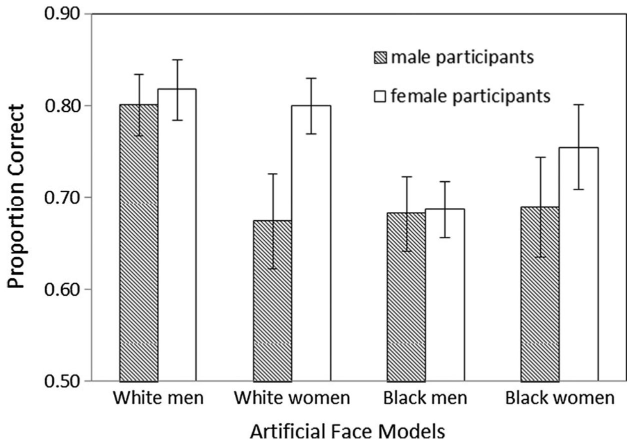

Fig. 2 Experiment 1: Identification of the "homosexual" face version (proportion correct) according to the gender and race of the stimuli (artificial face models) and the sex of the participants (chance level: 0.50). *Error bars* indicate ±SEM

analyses were carried out with participants (F_1) and items (F_2) as the random variables. The ANOVA revealed a significant Face Gender × Face Race interaction, and a significant Sex of Rater × Face Gender interaction (see Table 1). To improve the statistical study, data were organized in a long format (one observation per row) and submitted to a multilevel generalized mixed model. Mixed models combine F_1 and F_2 analysis treating both participants (raters) and items (faces) as random variables (e.g., see Baayen, Davidson, & Bates, 2008; Judd, Westfall, & Kenny, 2012; Westfall, Kenny, & Judd, 2014). In this case, the appropriate technique was a logistic regression analysis because the dependent variable is dichotomous. The output confirmed all previous ANOVA results: between the fixed effects, Face Race was significant (z of Wald = 17.03, $p < .0001$), and also the interactions of Face Race × Face Gender ($z = 20.58$, $p < .0001$) and Face Gender × Sex of Participants ($z = 5.54$, $p = .0186$).

Therefore, these data clearly showed that a sample of (young) Spanish people was able to choose correctly "homosexual" versions of the artificial faces created by manipulating the facial features identified as predictors of sexual orientation in a Canadian sample (Skorska et al., 2015). The results of the present experiment add perceptual validity and some cross-cultural consensus to the pattern of facial traits obtained in Skorska et al.'s study.

On the other hand, Black faces also yielded an above-chance score but certain race effect emerged since they were less well classified (−7%) than the White ones. Results suggest that, from a perceptual point of view, the seven facial parameters identified by Skorska et al. (2015) exhibit certain cross-race generalizability, although there was evidence of a race effect that will be discussed below.

Overall, the success of the women in the present experiment was 0.76, 95% CI [0.70, 0.82] and was 0.71 [0.63, 0.79] for men. Nevertheless, the Sex of Participants factor was not significant ($p = .149$) in the analysis of variance, likely because of the small number of male participants (indeed the effect size was $\eta_p^2 = .55$). The men performed the task with male faces better than with female faces (0.74, 95% CI [0.68, 0.80] vs. 0.68 [0.61, 0.75]) and, reciprocally, the women obtained better scores with female faces than with male faces (0.77 [0.72, 0.82] vs. 0.74 [0.69, 0.79]), resulting in the significant Face Gender × Sex of Participants interaction. This pattern was consistent with a certain own-sex bias reported in the face perception literature, especially for women. Evidence shows that women recognize and remember more female than male faces; in contrast, data about men are controversial (Rehnman, 2007).

Finally, the basis of the present experiment lies in a binary task in which participants had to choose between two faces which one they thought belonged to a person who was more likely to have a homosexual orientation. The results indicated that participants performed this task clearly above the chance level. In the next experiment, participants had to perform a more demanding perceptual task based on a rating scale. This task allows for a more continuous measure of the perceived sexual orientation in artificial face models.

Table 1 Significant effects (ANOVA) and effect sizes from Experiments 1 and 2

	Through subjects				Through items			
	F_1	*df*	*p*	η_p^2	F_2	*df*	*p*	η_p^2
Experiment 1								
Face Race	6.07	1,39	.018	.14	8.39	1,48	.006	.15
Face Race × Face Gender	7.53	1,39	.009	.16	11.59	1,48	.001	.15
Face Gender × Sex of Rater	4.11	1,39	.049	.09	6.10	1,48	.017	.11
Experiment 2								
Face Gender	32.20	1,38	<.001	.46	1.44	1,11	.256	.11
Face Version	53.75	4,152	<.001	.59	43.34	4,44	<.001	.80
Face Gender × Face Version	3.71	4,152	.007	.09	2.83	4,44	.036	.21

Partial eta-squared (η_p^2) refers to the proportion of variability in the dependent measure attributable to a factor. The effect size interpretations for η_p^2 values are as follows: .01 = small, .06 = medium, and .14 = large

Experiment 2

The second experiment studied how people judge on a rating scale the apparent sexual orientation of individual artificial faces created according to the main metrics identified by Skorska et al. (2015) as predictors of sexual orientation. This experiment was based on a more demanding task than a simple binary discrimination between two stimuli, and it was conducted to test whether people are sensitive to different degrees of intensity of the manipulated facial features and whether their judgments on sexual orientation are mainly continuous or categorical. Furthermore, as the faces were now presented one at time, the participants did not need to judge two versions of the same face in a single trial, which could consciously draw their attention to the manipulated features.

Methods

Participants

Forty-four young adults of both sexes participated in this experiment (31 females), whose age range was 19–29 years ($M = 20.75$; $SD = 2.62$). None had participated in Experiment 1. They were all undergraduates at the University Jaume I of Castellón (Spain), who volunteered in exchange for course credit. Of those who indicated ethnicity ($n = 35$), 86% were White/Caucasian and 4% were Hispanic/Latin American.

Materials

The experimental stimuli consisted of 120 artificial face models generated with the FaceGen Modeller 3.5 software (60 of White males and 60 of White females). All the facial models displayed the same default 3D position: yaw angle: 20.05° and pitch angle: 0.00°.

The method for creating the White male faces was as follows. First, the software generated 12 "neutral" faces randomly of European males. In a second step, from every "neutral" face, four additional faces were created by manipulating the same parameters as in Experiment 1: Cheeks—concave/convex, Nose—short/long, and Forehead—tilt forward/back. Concretely, two "homosexual" (gay) versions were obtained by adjusting the cheeks to +2 and +4 units, respectively, the nose to −1.5 and −3 units,[3] and the forehead to +2 and +4 units; two "heterosexual" versions were obtained by adjusting the same parameters: the first to −2 and −4 units, respectively, the second to +1.5 and +3 units, and the third to −2 and −4 units. In this way, it was possible to construct a set of five stimuli from each "neutral" face (by counting the own neutral version), with a range of equal steps from a "heterosexual" pole to a "homosexual" pole (see an example in Fig. 3). For the presentation, all 60 stimuli were mixed in random order.

The program created White female faces in a similar way. First, it generated 12 "neutral" faces randomly of European females. In a second step, from every "neutral" face, four additional faces were created by manipulating the same parameters as in Experiment 1: Nose—down/up, Mouth—protruding/retracted, Forehead—small/-large, and General gender control. Concretely, two "homosexual" (lesbian) versions were obtained by adjusting the first parameter to +1.5 and +3 units, respectively, the second to −1.5 and −3 units, the third to −1 and −2 units, and the general gender parameter 6 steps and 12 steps up; two "heterosexual" versions were obtained by adjusting the first parameter to −1.5 and −3 units, respectively, the second to +1.5 and +3 units, the third to +1 and +2 units, and the general gender parameter 6 steps and 12 steps down. Thus, as in male faces, it was possible to construct a set of five stimuli from each "neutral" female face (including the neutral version), which ranged in equal steps from a "heterosexual" pole to a "homosexual" pole (see an example in Fig. 3). For the presentation, all 60 stimuli were mixed in random order. White male and White female faces were presented in separated sessions.

Procedure

Participants saw all the generated faces in random order and in two separated counterbalanced blocks in a within-subjects design. The task was completed individually online through the university intranet (virtual classroom) in two sessions, with a rest lasting about 15 min between both.

Participants wrote their name and demographic data and received the following first instructions (in Spanish): "Recent research suggests that there may be subtle differences in facial structure associated with sexual orientation in both men and women. This is a perceptual study that extends that research. During each trial, an artificial face created by the computer will be displayed. Your task will consist in indicating, in your opinion, how likely you think that the person depicted has a homosexual orientation (i.e., toward the same sex). Here we use the term homosexual in a broad sense for both women and men." On another screen: "During each trial, the computer will display a male (female) face. You must mark on a scale from 1—No or very little probability of homosexual orientation to 7—Quite a high probability of homosexual orientation." During each trial, a computer displayed a single face with this label below: "Probability of homosexual orientation" on a scale of 1–7. Participants had to click on one of seven marks with a mouse, which ranged between the two extremes: 1 ("No or very little") and 7 ("very high").

For several unforeseen reasons, one woman did not complete the test of male faces, and three men did not complete the test of female faces. Rating scores were collapsed through participants and items within each gender set of stimuli. Thus for male faces, the rating means were calculated for every face version from the

[3] The size of the steps in each parameter (number of units) was adjusted to avoid greatly exaggerated versions on the extremes.

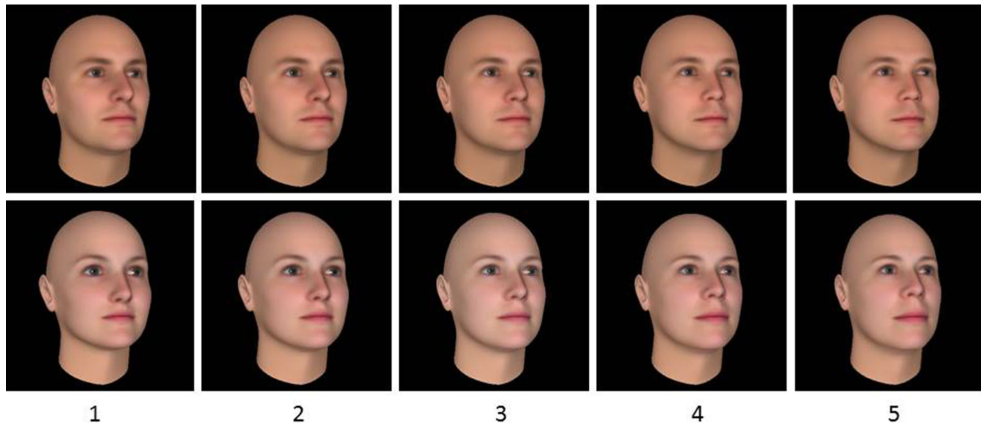

Fig. 3 Examples of the artificial faces created in Experiment 2, which ranged in equal steps from the more "heterosexual" (1) to the more "homosexual" (5) version. Male and female faces in the first and the second row, respectively. In each row, the central stimulus (3) is a "neutral" face generated randomly with the FaceGen software; the two left-hand faces (1, 2) are the "heterosexual" versions that derived from the neutral face, and the two right-hand faces (4, 5) are the "homosexual" versions that derived from the neutral face

more "heterosexual" version (1) to the more "homosexual" version (5). The same was done for female faces.

Results and Discussion

The data showed (Fig. 4) that homosexuality perceived from the artificial face models was continuous and almost a perfect linear function of the series of five artificial faces separated by equal steps when manipulating Skorska et al's. (2015) parameters.[4] For the male stimuli, the Pearson correlation between the rating scores and the series of face versions was $r = 0.99$ ($p = .001$), and for the female stimuli was $r = 0.98$ ($p = .003$). Overall, the female stimuli were perceived as being more homosexual than the male stimuli because the rating mean was larger (4.62 vs. 3.69) and the y-intercept coefficient from the regression equation was also higher (3.10 vs. 2.85). Sensitivity to the experimental manipulation of facial parameters while creating artificial models was also greater for the female than for the male stimuli (slope coefficients were 0.37 vs. 0.27). Thus, the perception in every step from one female face version to the next one was 0.37 points more homosexual on a scale of 1–7, whereas it was 0.28 points in the male faces. This difference in sensitivity was likely due to the manipulation of the general gender control in the FaceGen software, which corresponded to the fourth predictor in women identified in the multivariate analysis by Skorska et al. (2015).

A 2 (Sex of Rater) × 2 (Face Gender) × 5 (Face Version) mixed-model ANOVA was conducted. Separate analyses were carried out with participants (F_1) and items (F_2) as the random variables. The ANOVA revealed a significant Face Gender × Face Version interaction (see Table 1). As in Experiment 1, data were also organized in a long format (one observation per row) and submitted to a mixed model treating simultaneously subjects and items (faces) as random variables. Analysis followed Brysbaert's (2007) suggestions and basically corroborated the ANOVA results. The Face Gender × Face Version interaction resulted significant, $F(4, 4946) = 6.09$, $p = .00006$.

Figure 5 displays the mean ratings received for the male and female faces from men and women when segregating data according to Sex of Participants. The different pattern displayed by men and women when they rated the male models was striking (upper panel); men, unlike women, did not show sensitivity through the three central stimuli. Indeed a partial analysis confirmed that the rating scores given by men for male versions 2, 3, and 4 did not statistically differ from each other, unlike the scores by women, which differed significantly from each other. Another interesting observation was regarding the female face stimuli (lower panel): perceptively, there was hardly any difference between version 1 (an allegedly very heterosexual female face) and version 2 (a heterosexual female face), especially for men; that difference in women was marginally significant ($p = .097$). Nevertheless, this apparently differential pattern between male and female judgments should be viewed with caution, given the small number of male participants.

General Discussion

An issue with important social implications is whether sexual orientation can be accurately perceived by lay people. In this sense, a relevant source of information is the human face, although very

[4] Skorska et al. (2015) used the label "more puckered" mouth several times to refer to lesbian women; this facial trait corresponds to the FaceGen parameter called "Mouth protruding-retracted," which reached a partial correlation of −.42 with sexual orientation in women (Skorska et al., 2015, Table 2) (negative correlation indicates that lesbian women had less of that metrics; i.e., less retracted mouth). There is another FaceGen parameter called "Mouth–Lips puckered/retracted" which, in Skorska et al.'s work, yielded a lower partial correlation (−.31, Table 2). In the present study, the former one was the experimentally manipulated parameter.

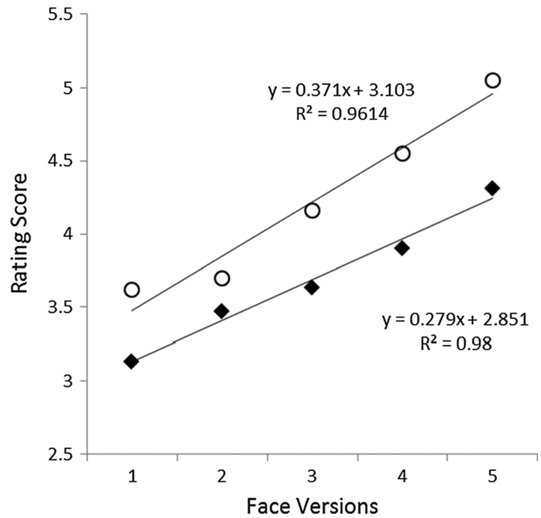

Fig. 4 Experiment 2: Regression lines and equations between the "homosexuality" rating scores and the five face versions, which ranged from the more "heterosexual" (1) to the more "homosexual" version (5). Data are separated for the male faces (*full diamonds*) ($r = 0.99$; $p = .001$) and the female faces (*open circles*) ($r = 0.98$; $p = .003$)

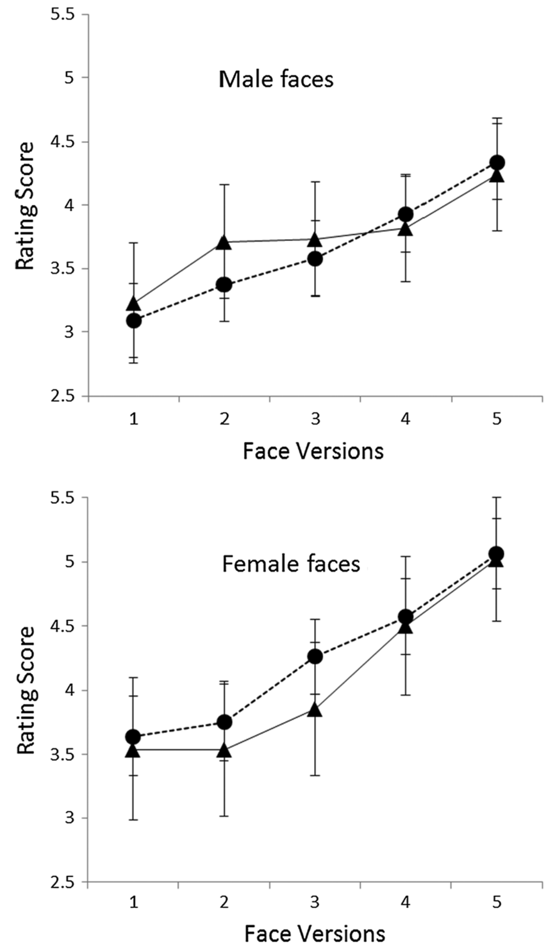

Fig. 5 Experiment 2: The "homosexuality" rating scores separated by the participants' sex (men: triangles; women: circles) according to the five face versions that ranged from the more "heterosexual" (1) version to the more "homosexual" version (5). The upper panel displays the data for the male face stimuli and the lower panel for the female face stimuli. *Error bars* indicate ±SEM

few studies have examined facial structure in relation to sexual orientation; indeed, the scientific literature includes only three studies of this kind (Hughes & Bremme, 2011; Skorska et al., 2015; Valentova et al., 2014).

The very extensive morphometric study by Skorska et al. (2015) identified three and four anatomical parameters as the more powerful predictors between gay/heterosexual men, and between lesbian/heterosexual women, respectively. An important step would be to test the perceptual validity of these metrics extracted from a Canadian sample. This was the main purpose of the present study, using artificial face models that varied solely in these specific metrics in order to examine their perception by (Spanish) people who belonged to a different geographical and cultural environment. The results of the two experiments presented herein showed a strong sexual orientation effect and that the anatomical cues identified by Skorska et al. actually affected participant's perceptions of sexual orientation. The basis of the first experiment was a binary discrimination task, and it extended the study scope by including Black artificial face models of both sexes. Overall, the data revealed a good accuracy level (0.74), which was well above chance level (0.50), and was 0.80 under the more favorable conditions (race/gender consistency between stimuli and participants). These figures are comparable and even higher than the data reviewed by Tskhay and Rule (2013) in their meta-analysis of research with stimuli based on real people. Tskhay and Rule found an overall correlation of $r = .29$ between perceived and actual (self-reported) sexual orientation. According to Rosenthal and Rubin's (1982) formula, the equivalent accuracy level is $50 + 50r = 64.5\%$. That is, the present artificial faces based on the selective manipulation of Skorska et al.'s anatomical predictors of sexual orientation were better classified than, in general terms, the stimuli—photographs, audio, video—obtained from real people (self-reported homo/heterosexual) as reviewed by Tskhay and Rule.

Interestingly, the responses to an open-ended question indicated that participants had issued their hetero/homosexual judgments based mainly on a holistic face representation, and guided by overall facial appearance, and that very few people realized or focused on the experimentally manipulated specific traits.

Data were sensitive to a race effect that could be accounted for different reasons. Black faces were less accurately classified than White faces, and this difference could be consistent with an own-group bias hypothesis. In the face perception field, research has repeatedly found that people generally better recognize and remember the faces that correspond to the same group as themselves from characteristics like age, race or sex (for a review, see Rehnman, 2007). The task of the first experiment did not involve recognizing or remembering an individual face, but carrying out a perceptual inference that likely implies a considerable cognitive load. Participants were better at discriminating "heterosexual" vs. "homosexual" artificial faces when the stimuli corresponded to the same race group (White faces). This is in line with previous studies which have demonstrated that Caucasian subjects recognize Caucasian faces more accurately than non-Caucasian faces (e.g., Tanaka, Kiefer, & Bukach, 2004). Nevertheless, we must be cautious with an own-group bias hypothesis because some previous research has not found a race effect in perceiving sexual orientation. Thus, Brambilla, Riva, and Rule (2013), Rule (2011), and Rule, Ishii, Ambady, Rosen, and Hallett (2011) did not observe an in-group race advantage for judging male sexual orientation.

Other possible explanations should be considered. First, our stimuli were based on manipulation of certain anatomical traits found by Skorska et al. (2015), but that study only examined facial structure characteristics in White participants. Thus, we do not know whether the same facial features would differ between Black gay/lesbian and heterosexual individuals, given the ethnic variation that is evident in facial structure. Second, some unnoticed interactions between race and sexual orientation may have influenced the judgments; for example, some of the facial features manipulated in the current study (e.g., degree to which the lips were puckered) interacted with facial features that differentiated White and Black people (e.g., Black people tend to have broader lips than White people). The same could be said about nose shapes. Furthermore, an important issue that should be considered is the evidence that racial stereotypes interact with gender phenotypes or stereotypes, affecting categorization of sex and sexual orientation. For example, Johnson, Freeman, and Pauker (2012) observed that sex categorization varied systematically as a function of race: Black faces were associated with male stereotypes, whereas Asian faces were associated with female stereotypes. Consistent with that bias, Johnson and Ghavami (2011) found that Black men were more likely to be rated as heterosexual because being Black is associated with masculinity. Thus, our results could be influenced by the interaction between both types of stereotypes associated with race and gender. In summary, several explanations are possible to account for the higher score obtained from White faces in the present study.

On the other hand, the Face Gender × Sex of Participants interaction was significant in the first experiment because participants displayed better performance with faces of their own gender. This was especially true for women, who were clearly better than men at detecting homosexuality in the female faces in both the Black and White faces. This women's superiority with female faces is in line with a general superiority of women in face recognition. Accumulative evidence has shown that women of different ages recognize more faces than men and that they are particularly efficient at recognizing female faces (Rehnman, 2007).

Beyond a simple binary choice between two faces, the second experiment was more perceptually demanding and prompted participants to rate on a scale the "homosexuality" probability of a series of artificial faces created by varying Skorska et al.'s key parameters in the steps with the same value. Once randomly mixed, the faces were individually presented and the rating scores displayed an almost perfect linear regression according to the parameter steps. Once again, some differences associated with both the participant's sex and face gender emerged. Overall, ratings of female faces were more sensitive to experimental manipulation than male faces, likely because of the manipulation of the FaceGen software's general gender control, which could exert a stronger effect on the whole face appearance. When examining the data separated by sex of raters, men showed a different sensitivity pattern, this time for male faces: unlike women, men did not show sensitivity through the three central stimuli and did not find them different in homosexuality probability; for men, the five face versions perceptually became only three: the first one, the three central stimuli in a same pool, and the fifth one. Conversely, women were sensitive to the five versions and their rating scores statistically differed from all the other stimuli, which is once again in line with women's proven face recognition superiority.

In summary, the present study supported the perceptual validity of the seven multivariate predictors of sexual orientation identified by Skorska et al. (2015) and opens up new avenues to further research this issue by experimentally manipulating artificial face models. Future work should go more deeply into the understanding of a possible own-race effect, including non-Caucasian participants, and should also replicate gender differences with larger male samples. Beyond the seven key parameters studied herein, Skorska et al. identified 11 and 17 facial features at the univariate level, which differed between gay/heterosexual men and lesbian/heterosexual women, respectively. Further experiments could help us examine the perceptual validity of these facial parameters and their relative weights. Finally, it is important to conduct extensive morphometric studies in other cultural environments to establish cross-cultural comparisons and to test the hypothetical universality or certain subtle facial features associated with sexual orientation.

Compliance with Ethical Standards

Conflict of interest None.

Informed Consent All individual participants included in this study gave their informed consent.

References

Ambady, N., Hallahan, M., & Conner, B. (1999). Accuracy of judgments of sexual orientation from thin slices of behavior. *Journal of Personality and Social Psychology, 77,* 538–547.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412. doi:10.1016/j.jml.2007.12.005.

Berger, G., Hank, L., Rauzi, T., & Simkins, L. (1987). Detection of sexual orientation by heterosexuals and homosexuals. *Journal of Homosexuality, 13,* 83–100.

Brambilla, M., Riva, P., & Rule, N. O. (2013). Familiarity increases the accuracy of categorizing male sexual orientation. *Personality and Individual Differences, 55,* 193–195. doi:10.1016/j.paid.2013.02.023.

Brysbaert, M. (2007). *"The language-as-fixed-effect fallacy": Some simple SPSS solutions to a complex problem (Version 2.0).* Royal Holloway: University of London.

DeBruine, L., Jones, B., Unger, L., Little, A., & Feinberg, D. (2007). Dissociating averageness and attractiveness: Attractive faces are not always average. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1420–1430. doi:10.1037/0096-1523.33.6.1420.

Ferreira, L. (2015). *How many African states outlaw same-sex relations?* (*At least* 34). Report retrieved 28 August 2015 from https://africacheck.org/.

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.

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.

Johnson, K. L., Freeman, J. B., & Pauker, K. (2012). Race is gendered: How covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology, 102,* 116–131.

Johnson, K. L., & Ghavami, N. (2011). At the crossroads of conspicuous and concealable: What race categories communicate about sexual orientation. *PLoS ONE, 6,* e18025. doi:10.1371/journal.pone.0018025.

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103,* 54–69. doi:10.1037/a0028347.

Lefevre, C., Ewbank, M., Calder, A., von dem Hagen, E., & Perrett, D. (2013). It is all in the face: Carotenoid skin coloration loses attractiveness outside the face. *Biology Letters, 9*(6), 20130633. doi:10.1098/rsbl.2013.0633.

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

Rieger, G., Linsenmeier, J. A. W., Gygax, L., García, S., & Bailey, J. M. (2010). Dissecting "gaydar": Accuracy and the role of masculinity-femininity. *Archives of Sexual Behavior, 39,* 124–140. doi:10.1007/s10508-008-9405-2.

Rosenthal, R., & Rubin, D. B. (1982). A simple general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74,* 708–712. doi:10.1037/0022-0663.74.2.166.

Rule, N. O. (2011). The influence of target and perceived race in the categorisation of male sexual orientation. *Perception, 40,* 830–839. doi:10.1068/p7001.

Rule, N. O. (2017). Perceptions of sexual orientation from minimal cues. *Archives of Sexual Behavior*, *46*, 129–139. doi:10.1007/s10508-016-0779-2.

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., Ambady, N., & Hallett, K. C. (2009a). Female sexual orientation is perceived accurately, rapidly, and automatically from the face and its features. *Journal of Experimental Social Psychology, 45,* 1245–1251. doi:10.1016/j.jesp.2009.07.010.

Rule, N. O., Ishii, K., Ambady, N., Rosen, K. S., & Hallett, K. C. (2011). Found in translation: Cross-cultural consensus in the accurate categorization of male sexual orientation. *Personality and Social Psychology Bulletin, 37,* 1449–1507. doi:10.1177/0146167211415630.

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

Singular Inversions. (2010). *FaceGen Modeller*, Version 3.5. www.facegen.com.

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.

Sylva, D., Rieger, G., Linsenmeier, J. A. W., & Bailey, J. M. (2010). Concealment of sexual orientation. *Archives of Sexual Behavior, 39,* 141–152. doi:10.1007/s10508-008-9466-2.

Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition, 93,* B1–B9. doi:10.1016/j.cognition.2003.09.011.

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.

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General, 143,* 2020–2045. doi:10.1037/xge0000014.