# “I SEE GAY PEOPLE” GAYDAR ABILITIES IN A REAL-WORLD DISTRIBUTION

By

Benjamin P Skillman

A Thesis Presented to

The Faculty of Humboldt State University

In Partial Fulfillment of the Requirements for the Degree

Master of Arts in Psychology: Academic Research

Committee Membership

Dr. Amanda C. Hahn, Committee Chair

Dr. Christopher L. Aberson, Committee Member

Dr. Amber M. Gaffney, Committee Member

Dr. Christopher L. Aberson, Program Graduate Coordinator

July 2019

## Abstract

"I SEE GAY PEOPLE" GAYDAR ABILITIES IN A REAL-WORLD DISTRIBUTION

Benjamin P Skillman

"Gaydar" is the colloquial term for identifying someone's sexual orientation from physical cues. Past literature has shown that people can identify someone's sexual orientation at above chance levels. Past literature has used a 50/50 split of gay and straight faces and used non-standardized images, which can induce confounds in the results. The present study examined gaydar accuracy in a realistic distribution of straight and gay faces using standardized images and examined facial morphology for differences between gay and straight men. Participants were not found to have above chance accuracy for identifying gay faces. PCA did not identify reliable shape differences between gay and straight men's faces. Participants past contact with gay men did not affect their gaydar accuracy. When examining perceived sexual orientation of the faces, faces that were rated as more masculine by a separate sample tended to be rated as gay less often than the feminine faces by those in the rating task. The finding that femininity is associated with being perceived as gay is supported by previous literature. The current study calls into question the idea that people have accurate gaydar abilities and the idea that there are reliable facial differences between gay and straight men.

**Acknowledgements**

I want to thank Dr. Valentova for graciously sharing her stimuli with me, I would not have been able to finish without them. I want to thank all the research assistants who helped me collect data and map faces. I want to thank Dr. Hahn for her immense amount of help long the way. Most importantly, I want to thank my mom for always being my biggest supporter.

## Table of Contents

## List of Tables

## List of Figures

## Introduction

Every day, people make judgements about those around them based on physical traits. The notion that we can make accurate perceptions about individuals from brief interactions alone is often referred to as the "kernel of truth hypothesis" (e.g., Berry, 1990; Penton-Voak, Pound, & Perrett, 2006; Todorov, Said, Engell, & Oosterhof, 2008). In their meta-analytic review of this literature, Ambady and Rosenthal (1992) summarized studies investigating the accuracy of people's judgements from these thin slices of behavior. Their analysis suggested that people are, indeed, accurately able to judge a variety of personality traits from faces, voices, and/or bodies based on interactions as short as 30 seconds or as long as five minutes. In particular, the traits most accurately assessed included: teaching efficacy, existence of deception, and patient wellbeing. Additional research has shown that we can accurately perceive additional traits such as threat (Oosterhof & Todorov, 2008), trust, and dominance (Fruhen, Watkins, & Jones, 2015).

Whereas previous literature has suggested that there may indeed be a "kernel of truth" in these perceptual judgements, this work has primarily focused on potentially facultative personality traits. More recently researchers have begun to investigate whether these perceptions extend to aspects of identity, such as sexual orientation. The concept of being able to identify someone's sexual orientation from visual cues is colloquially known as "gaydar". In an early study on this topic, participants were found to be able to accurately identify a target's sexual orientation from 10 second and one second silent

videos, and from six stills shots featuring the participant's whole body (accuracy rates approx 62%; Ambady, Hallahan, Conner, 1999). A recent study that sparked a great deal of controversy in the media suggested that it may even be possible to train AI to detect sexual orientation from photographs and suggested that AI was more accurate than people at judging sexual orientation from pictures of faces (Wang & Kosinski, 2018).

The idea that people can accurately judge personality characteristics from facial cues has led to the notion that people with different personality traits may have underlying differences in facial morphology (Oosterhof & Todorov, 2008). Given that previous research has suggested people may be able to make accurate inferences regarding sexual orientation from facial cues, is it possible that there are detectable phenotypic differences between straight and gay individuals? Studies directly examining facial morphology have shown that, compared to straight men, gay men tend to have less symmetrical faces (Hughes & Bremme, 2011) and more feminine faces (Skorska, Geniole, Vrysen, McCormick, & Bogaert, 2015; Robertson, Kingsley, & Ford, 2019). More specifically, gay men tend to have more rounded and larger chins, smaller noses, eyes that are closer together, and mouths with more downward oriented corners (Valentova, Kleisner, Havlíček, & Neustupa, 2014). One potential mechanism for the development of morphological differences between gay and straight men is proposed differing levels of testosterone exposure in prenatal development (Androgen Signaling Theory; Rice, Friberg, Gavrilets, 2012). This is supported by the finding that gay and straight men also show differences in sexually dimorphic brain regions such as the hypothalamus, which is rich in androgen receptors (Levay, 1991). Furthermore,

testosterone affects skeletal growth and gay men have been found to be shorter and have shorter long bones than straight men (Skorska & Bogaert, 2017; Martin & Nguyen 2004).

When it comes to people's ability to perceive sexual orientation, Rule and Ambady (2008) found that sexual orientation could be reliably discerned at above chance levels just from looking at a man's face for as little as 50 milliseconds. When judging sexual orientation of women, sexual orientation could be accurately categorized from only 40 milliseconds (Rule, Ambady, & Hallett, 2009). Although some research has suggested that perceptions of sexual orientation involve holistic processing of faces (Tabak & Zayas, 2012), several studies indicate that the entirety of a person's face is not necessary to make accurate judgements about a person's sexuality; people are more accurate than chance at categorizing men by looking only at their hair, eyes, or mouth alone (Rule, Ambady, Adams, & Macrae, 2008). This research highlights the influence of contextual cues in sexual orientation perception, in addition to potential morphological differences in the faces of gay and straight men.

Gaydar abilities appear to vary both between and within individuals. For example, individuals who report having lower prejudice toward gay people have higher accuracy of sexual orientation judgements when judgments are made from unstandardized, public-domain images (Rule, Tskhay, Brambilla, Riva, Andrzejewski, & Krendl, 2015). Women's gaydar abilities fluctuate across the menstrual cycle, with more accurate judgments during peak ovulation when judging men's but not of women's faces, and are sensitive to context, with more accurate judgments being made when women are

primed with mating/reproduction cues (Rule, Rosen, Slepian, & Ambady, 2011). Additionally, when individuals make a judgment regarding the sexual orientation of a target, the speed of that judgment has been shown to affect accuracy - with faster judgments being more accurate (Rule et al., 2009).

In addition to these within and between individual variations in gaydar abilities, research has also suggested that there may be a learned or experiential component to the ability to accurately perceive sexual orientation from an individual's physical appearance. A recent study found that increased contact or familiarity with gay people was related to increased gaydar accuracy (Brambilla, Riva, & Rule, 2013). The notion that experience or expertise with a class of faces improves perception and/or recognition abilities is in line with decades of work on the Other Race Effect, which refers to the robust finding that faces of one's own race are easier to recognize and perceive than faces of other races (Goldstein & Chance, 1985; Lindsay, Jack, & Christian, 1991; Walker & Tanaka 2003). Early research on the impact of experience/contact on the Other Race Effect demonstrated that participants trained to perceive Japanese faces showed better recognition of Japanese faces than those who were not trained to perceive Japanese faces (Goldstein & Chance, 1985). Additionally, Asian participants attending a majority white Canadian school showed increased processing for both Caucasian and Asian faces while Caucasian students showed increased processing for only Caucasian faces (Tanaka, Kiefer, & Bukach, 2004). Together, this research demonstrates that contact or experience is a key factor in how faces are processed and recognized, which suggests that contact or familiarity with gay people may affect how accurately their faces are processed.

Though people tend to make more accurate judgements about the sexual orientation of individuals from their own race/culture (Valentova, Rieger, Havlíček, Linsenmeier, & Bailey, 2011), evidence suggests they may still be able to make accurate assessments of the sexual orientation of people from other races/cultures (Rule, Ishii, Ambady, Rosen, & Hallett, 2011; Valentova et al., 2011). Though gay stereotypes may vary from culture to culture, the idea of what a gay person "looks like" may not vary as widely between cultures. Overall gender congruence (sex typicality) appears to be an important cue for perceptions of an individual's sexual orientation. Many studies show that people rely on the masculinity/femininity of the target when making their judgements about a person's sexual orientation (Freeman, Johnson, Ambady, & Rule, 2010; Lyons, Lynch, Brewer & Bruno, 2014; Rieger, Linsenmeier, Gygax, Garcia, & Bailey, 2010; Valentova et al., 2011). For example, feminine men are more likely to be rated as gay than straight and masculine women are more likely to be rated as lesbian than straight. This appears to extend to different modalities. For example, increased perceived femininity in men's voices is linked to increased attributions of homosexuality (Valentova & Havlíček, 2013). Additionally, research has suggested these perceptions occur from bodily movement as well. One study found that participants more frequently judged individuals walking on a treadmill as gay if they had a more female-typical movement, even with a masculine body shape (Johnson, Gill, Reichman, & Tassinary, 2007).

**Problems with Gaydar Research**

Research on face-based gaydar has several confounds that affect much of the current literature. First, many of the facial stimuli used in the past literature are non-standardized photos, primarily composed of pictures gathered from dating or social media websites. These pictures can vary on quality, photographic conditions, expression of the target, and the degree of additional contextual cues to sexual orientation (e.g., environment displayed in background, hairstyle and clothing/accessories, etc.). Recent studies investigating social perceptions from facial cues have demonstrated that variation in photographic conditions (e.g., lighting, focal length, camera-to-head distance, position/posture of subject) can distort the morphometrics of the skull (Eliášová & Krsek, 2007) and affect assessments of masculinity/femininity (Třebický, Fialová, Kleisner, & Havlíček, 2016), dominance (Hehman, Leitner, & Gaertner, 2013), as well as health (Stephen, Smith, Stirrat, & Perrett, 2009). While the use of non-standardized photos from dating websites may add to the external validity of the study, it decreases the internal validity. Gay men tend to have higher quality pictures than straight men (Cox, Devine, Bischmann, & Hyde, 2016) which could be artificial cue to sexual orientation. When this difference was controlled for, people relied much more heavily on stereotypic statements and very weakly on the picture of the face when making judgements about sexual orientation. Rule, Johnson, and Freeman (2016) investigated the quality of the picture sets used in past studies and found significant quality differences in five of their 13 stimulus sets. When standardized photos are used, the accuracy of sexual orientation

judgements tend to fall below chance (Valentova & Havlíček, 2013), calling into question the notion that it is possible to accurately assess sexual orientation from potential differences in facial morphology between gay and straight men alone.

Beyond picture quality, using photos from dating websites presents the issue of self-presentation bias; people are unlikely to portray how they look on a day to day basis in their profile picture (Toma & Hancock, 2010). Some individuals will use makeup or better lighting to make their face look different than they do normally. One study found that many people go so far as having a glamour shot taken to use as a dating profile picture (Witty, 2008). In an attempt to control for self-presentation bias, the researchers used pictures that the target was "tagged" in on Facebook (Rule & Ambady, 2008). However, this method may not actually prevent self-presentation bias; the person may not have posted the picture, but it is entirely likely that they still had a say in how they looked and whether or not the photo was uploaded.

Another flaw in the facial gaydar literature is the distribution of facial stimuli used in studies. Most, if not all, of the studies examining gaydar have used a 50/50 distribution of pictures of gay and straight men (e.g., Rule & Ambady, 2008, Cox et al., 2016), but more often than not, LGBT people will be the minority group in a social environment (Gates, 2017). Past literature states that around 5% of the US population is gay (Plöderl, 2014). A recent *Gallup Poll* states that the population of LGBT people in the US at 4.1% (Gates, 2017). In the UK, the Office for National Statistics' report shows the LGB population at 2%, but 4.1% for those 16-24 (Knipe, 2017). Because homophobia is still somewhat prevalent in the US and the UK (e.g. Drydakis 2015), gay

individuals are likely underreporting or may not feel comfortable reporting their sexual orientation, suggesting that the actual LGBT population may be closer to 5% of the total population. Cox et al. (2016) recently pointed out the poor generalizability of studies using these non-representative distributions. Half-and-half distributions may work for investigating differences between groups and the mechanisms of these judgements but they cannot generalize to real world accuracy. Most gaydar studies report an overall accuracy rate between 50% and 70% for these 50/50 distributions, however, it is unclear if this above chance accuracy would be found in samples that matched a real-world distribution of gay and straight individuals. This idea is backed by the ignoring of the base-rate fallacy in most, if not all, studies on gaydar (Plöderl, 2014). The base-rate fallacy is when one ignores general information (the base-rate) and uses information for specific cases when making decisions (Cox et al., 2016). In the gaydar field, researchers are using a manufactured distribution of 50/50. Because these proportions don't match the real world, the accuracy ratings are not only generalizable, they are grossly inflated. Plöderl (2014) goes so far as to suggest that the interpretations of the past studies are "unethical".

Gaydar judgements may have negative repercussions. People who were not morally opposed to prejudice but still wanted to be viewed as non-prejudiced tended to give more intense shocks to a person who was implied to be gay than a person who was explicitly labeled as gay (Cox & Devine, 2014). Because the imaginary subject was not explicitly gay, participants could act prejudiced without being viewed as prejudiced. When looking at the control conditions, shock intensity was unrelated to internal and

external motivation to respond without prejudice when the imaginary confederate was straight. That being said, people seem to have fallible opinions of gaydar. When people are told that gaydar is just stereotyping, and therefore not real, they tend to ignore the content of stereotypical statements when making judgements of sexual orientation from a face (Cox et al., 2016).

**Predictions**

More recently, research has focused on the caveats of sexual orientation identification. The literature is divided on whether or not face-based gaydar is the detection of underlying morphological (i.e. physical) differences in the faces of straight and gay men or whether it is simply stereotyping based off contextual or stylized cues like hair, interests, or expression. The purpose of my study is to investigate if the removal of contextual cues in facial photographs effects people's gaydar abilities.

The aims of this study are: to investigate if above chance accuracy of judgments of sexual orientation can be found in a realistic distribution of gay and straight faces while controlling for picture quality, to investigate the impact of contact/familiarity on the accuracy of these judgments, and to investigate the role of potential morphological differences in the face that contribute to the perception of sexual orientation. The facial photographs will be analyzed for morphological differences using facial mapping software to determine which aspects of facial appearance contribute to perceptual judgments of sexual orientation. I predict that people will show above chance accuracy for identifying the gay and straight faces and the gay faces alone, those with increased

contact with gay men will show increased gaydar accuracy, and gay men and straight men will show differences in face shape.

These hypotheses, methods, and planned analyses were pre-registered at https://osf.io/fqd64/. Some minor modifications were made to the planned analyses in light of newly established protocols in the field.

## Method

### Participants

Participants ($N$ = 97) were brought into the lab to complete the study on Qualtrics. The sample consisted mainly of Humboldt State students recruited from SONA ($M_{age}$ = 21.22, $SD_{age}$ = 4.40). The sample consisted of mainly women (n = 67), some men (n = 14) and, some non-binary (n = 9) participants. Sample size was derived from an a priori power analysis. Participants enrolled in eligible courses received extra credit.

### Facial Stimuli

Ninety-six full-face photographs (52 straight) of adult Czech men were obtained from a collaborator for use in this study. All photographs were taken under standardized photographic conditions in two separate sets. Six of these images were excluded due to being duplicates across both sets and 14 were excluded after visual inspection due to suboptimal standardization (e.g., visible hair that could not be masked out). A total of 76 images (38 straight, 38 gay) were used in the final perceptual task. The photographed

men were recruited through website advertisements, pamphlets distributed at local bars, through a radio broadcast and snowball sampling. All individuals photographed were between 18 and 35 years old. Photographs were taken in a small windowless room against a constant background, under standardized lighting conditions, and participants were instructed to pose with a neutral expression. Camera-to-head distance (1.5m) and camera settings were held constant for all photographs. Photographs were taken using a Canon 350D camera with the focus Canon EF 50/1.8 II (see Valentova & Havlíček 2013).

For use in the study, the facial photographs were mapped using a standard 189-point template that demarcates landmark and semi-landmark features using Webmorph (DeBruine, 2017). Images where then aligned based on interpupillary distance and masked using a white background, thus removing the hairstyle and other potential contextual cues to sexual orientation. The men in these photographs reported demographic information, including sexual orientation. Sexual orientation was assessed using the 7-point Kinsey Scale (Kinsey, 1953). For the purposes of the experimental task, these men were categorized as straight if they report a score of 0-2 on the Kinsey Scale, gay if they report a score of 4-6 on the Kinsey Scale. Men reporting bisexual orientation (i.e., a score of 3) were not included in the image set.

**Procedure**

Participants completed a two-alternative forced choice judgement for each of the aforementioned stimuli in the lab, whereby they indicated if they thought the individual

was straight or gay. Participants were randomly assigned to one of two conditions: a standard experimental condition which consisted of a 50:50 distribution of gay:straight faces (76 faces presented in total, 38 gay, 38 straight), or a real-world condition which consisted of a 7:93 ratio of gay:straight faces (40 faces presented in total, 3 gay, 37 straight - note that the total number of faces included in this condition was limited by the number of faces available in the imageset; Gates, 2017). A control condition with 95:5 gay:straight faces was originally proposed but was dropped due to a coding error that would have made the ratios between this and the real-world condition unequal.
The three gay faces were randomly selected from the total available set of 38 gay faces and the identities presented varied between participants to ensure that perceptual abilities weren't affected by any one individual in the imageset. All faces were presented in a fully randomized order for both conditions.

Following the rating task, participants reported their confidence in their "gaydar" abilities using a 5-point scale, where 1 (*"Not accurate at all"*) and 5 (*"Extremely accurate"*; following Brambilla et al., 2013). Specifically, they were asked "How accurate do you think your gaydar abilities are?" Participants also reported their amount of contact with gay men in the past using the same five questions used in Brambilla et al. (2013).

A separate group of 20 research assistants from research labs at Humboldt State University rated the men's faces on masculinity/femininity on a seven point scale from 1 (*Very Feminine*) to 7 (*Very Masculine*) , see Figure 1 for distribution of these ratings for the gay and straight faces. Higher scores for perceived masculinity/femininity indicate

more masculine faces. The utilization of a separate group avoided participants being influenced by their judgements of the target's sexual orientation.


*Figure 1.* Rating of perceived masculinity/femininity across sexual orientation of the targets

## Results

Overall accuracy was computed for each participant by dividing the number of faces for which sexual orientation (either gay or straight) was correctly identified by the total number of faces seen ($M$ = 53.83%, $SD$ = 12.37%). Accuracy for the perception of gay faces specifically was computed for each participant by dividing the number of gay faces correctly identified by the total number of gay faces seen ($M$ = 43.23%, $SD$ = 24.89%). A contact score for each participant was computed by averaging the five contact questions from Brambilla et al. (2013), ($\alpha$ = .81). Overall, reported confidence in gaydar abilities appeared fairly low ($M$ = 2.18, $SD$ = 0.86); none of the participants indicated the highest confidence (*"Extremely accurate"*).

### Hypothesis 1

To investigate if participants (irrespective of the condition they were assigned to) were accurate at perceiving the sexual orientation of faces at above chance levels, a one sample t-test was used. Since the task was dichotomous, 50% accuracy represents chance. Accuracy scores were arcsine transformed to ensure compliance with assumptions, t-test statistics represent the arcsine transformed scores while the means and standard deviations are untransformed for interpretability. Overall, participants performed at above chance levels, $t(82) = 2.86$, $p < .01$, $d = 0.31$, $BF_{10} = 5.34$, for overall accuracy of the detection of sexual orientation from standardized faces. When looking at the accuracy of gay faces only, participants did not perform significantly different than

chance, $t(82) = -1.57$, $p = .12$, $d = 0.17$, $BF_{10} = 0.39$, indicating a tendency to over perceive gay men's faces as being straight.

A Welch t-test was used to compare the performance of participants in the real-world and the experimental conditions. Participants in the real-world condition ($M$ = 52.00%, $SD$ = 29.48%) were significantly better at identifying gay faces than participants in the experimental condition ($M$ = 34.77%, $SD$ = 14.28%), $t(56.08) = -3.05$, $p < .01$, $d$ = 0.66, $BF_{10} = 7.96 \times 10^{19}$.

**Hypothesis 2**

To investigate if those with more contact were more accurate than those with less contact, a regression was conducted with accuracy of categorization of gay faces as the dependent variable and condition, amount of contact, and confidence of gaydar abilities as predictors. Contact, confidence, and accuracy of gay faces were assessed for normality using 99% confidence interval of skew and kurtosis. Accuracy of categorization of gay faces was positively skewed and transformed to normality with a square root transformation. Transformations did not affect the outcome of the analysis so untransformed results are presented for interoperability. Predictors were also centered. While the overall model was significant, $F(7, 89) = 2.17$, $= .04$, $R^2$, $= .15$., only one of the predictor variables was a significant predictor. As shown in the t-test, those in the real-world condition tended to perform better than those in the experimental condition. Contact, confidence, and/or any of the interactions among these variables were not significant predictors of accuracy of gay faces (See Table 1). Figure 2 displays

the relationship between condition and confidence and Figure 3 displays the relationship between condition and contact. A robust regression was used to ensure compliance with assumptions. The robust regression did not provide different results.

*Table 1*. Regression Statistics for Condition, Contact, & Confidence for Accuracy of Gay Faces

| Variable | b* | *p* | $sr^2$ |
|---|---|---|---|
| Condition (Real-World) | .31 | .01 | .09 |
| Contact | .02 | .88 | .00 |
| Confidence | -.02 | .88 | .00 |
| Condition x Contact | .02 | .84 | .00 |
| Condition x Confidence | -.02 | .85 | .00 |
| Contact x Confidence | -.02 | .86 | .00 |
| Condition x Contact x Confidence | .12 | .35 | .01 |

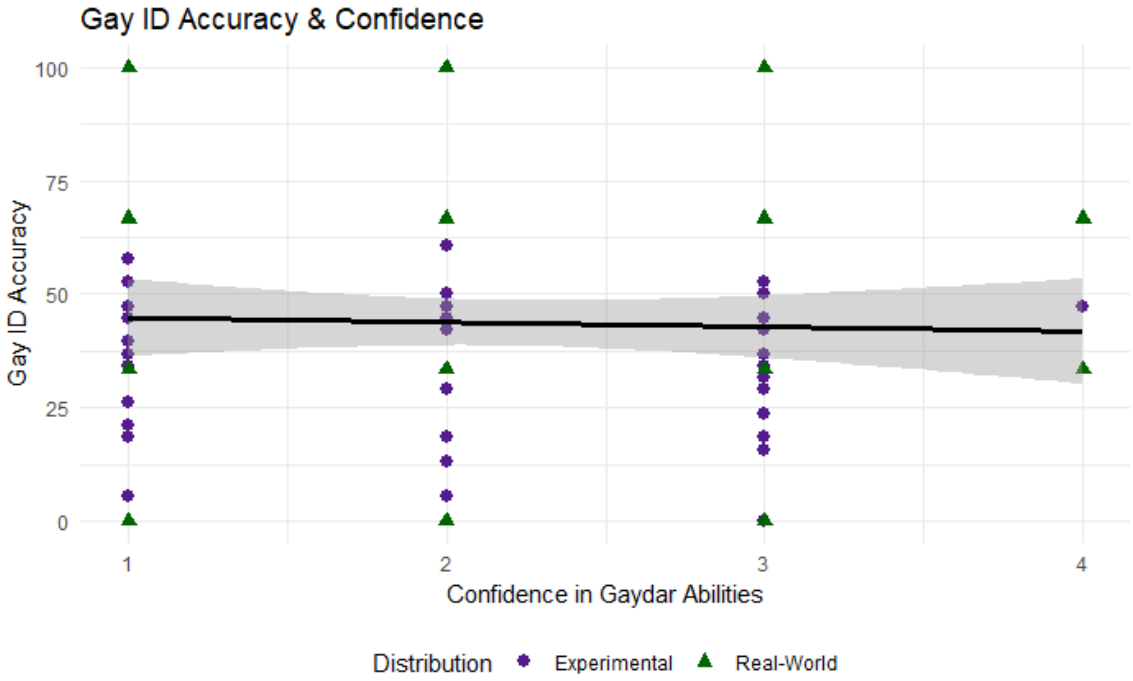


*Figure 2.* Confidence of accuracy and accuracy of identifying gay faces

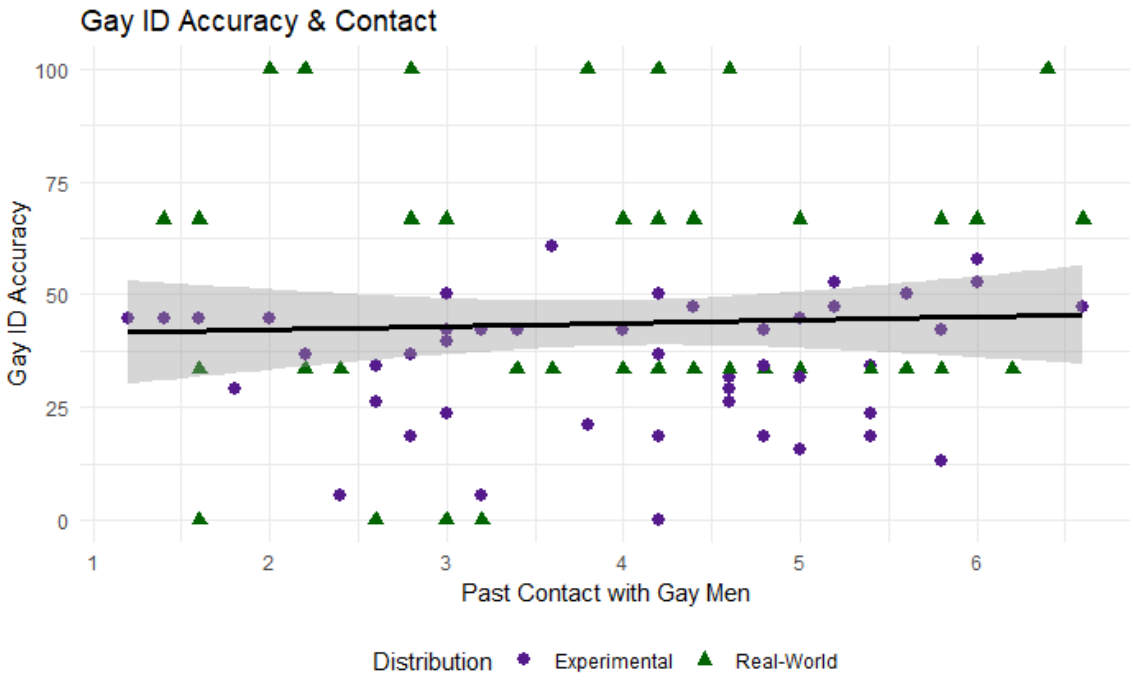


*Figure 3.* Past contact with gay men and accuracy of identifying gay faces

**Hypothesis 3**

**Calculating a Sexual Orientation Vector.** A principal components analysis (PCA) was conducted on the mapped facial points using publicly available code from https://osf.io/98qf4/. The first six principal components, explaining a total of 72% of variance in face shape were selected according to the Broken Stick Model. The first principal component was selected according to the Wilk's Lambda criterion for a linear discriminant analysis (LDA). The scores were used to create a sexual orientation vector with gay and straight prototypes at either end. Each face can then be given a sexual orientation vector score based on its respective position along the vector, with higher scores indicating a more gay typical face shape, and lower scores indicating a more straight typical face shape. This method of analyzing facial morphology follows previously established protocols (Holzleitner et al. 2018; Hahn et al. 2018).

**Calculating Facial Masculinity/Femininity.** From the 20 masculinity/femininity ratings of the 76 faces, two of the faces had a missing score from one rater. The mean rating for the face was imputed in place of missing rating and taken as part of the overall average score for the face. Refer to Figure 1 for a visual representation of the distribution of these masculinity/femininity ratings for the gay and straight faces.

**Calculating Perceived Sexual Orientation.** A perceived sexual orientation score was computed for each face by dividing the number of times a face was labeled as gay by the total number of times the face was seen by participants in the gaydar task. Therefore,

higher scores indicate that a face was perceived as gay more often than straight (regardless of actual sexual orientation).

**Statistical Analyses.** Perceived masculinity/femininity, perceived sexual orientation, and the sexual orientation vector were assessed for normality. Perceived sexual orientation and the sexual orientation vector were positively skewed and therefore transformed to normality using a square root transformation. Transformations did not influence significance of results so untransformed results are reported for interoperability.

The sexual orientation vector and perceived masculinity were entered into a logistic regression predicting *actual* sexual orientation. Unsurprisingly, the sexual orientation vector significantly predicted actual sexual orientation, OR = 1.31, *OR 95% CI* [1.03, 1.72], $p = .04$, where higher scores on the vector related to a greater likelihood of the target self-reporting as gay. Perceived masculinity/femininity did not predict actual sexual orientation, OR = 0.77, *OR 95% CI [0.35, 1.65]*, $p = .50$.

The sexual orientation vector and perceived masculinity/femininity were next used to predict *perceived* sexual orientation, $F$ (2, 73) = 17.47, $p < .001$, $R^2 = .32$. The sexual orientation vector did not significantly predict perceived sexual orientation, $b^* = -0.13$, $p = .186$, squared $sr^2 = .02$ (see Figure 4). Perceived masculinity/femininity was a significant predictor whereby more masculine faces (higher scores) were rated as gay less often than feminine faces, $b^* = -0.53$, $p < .001$, $sr^2 = 0.26$. Figure 5 shows the relationship between perceived masculinity/femininity and perceived sexual orientation. Upon graphing the data, a potential outlier emerged. The exclusion of the potential outlier did not affect the significance of results, the outlier was not removed in

the presented statistics. A robust regression was used to ensure compliance with assumptions. The robust regression did not provide different results.

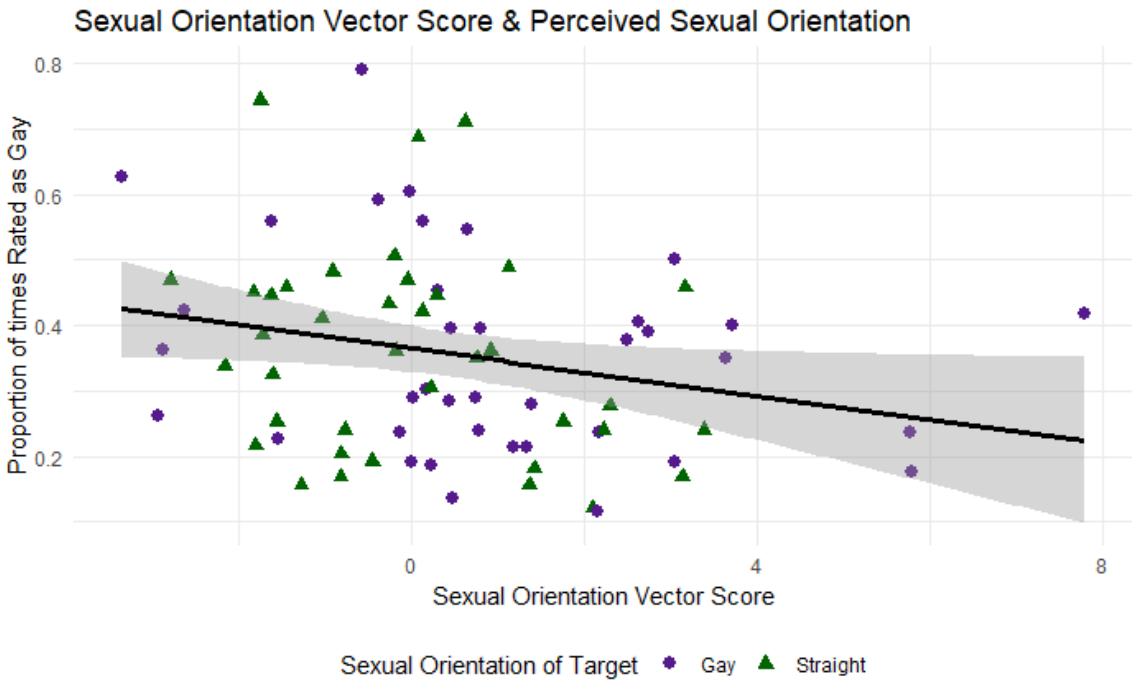


*Figure 4.* Sexual orientation vector score and perceived sexual orientation

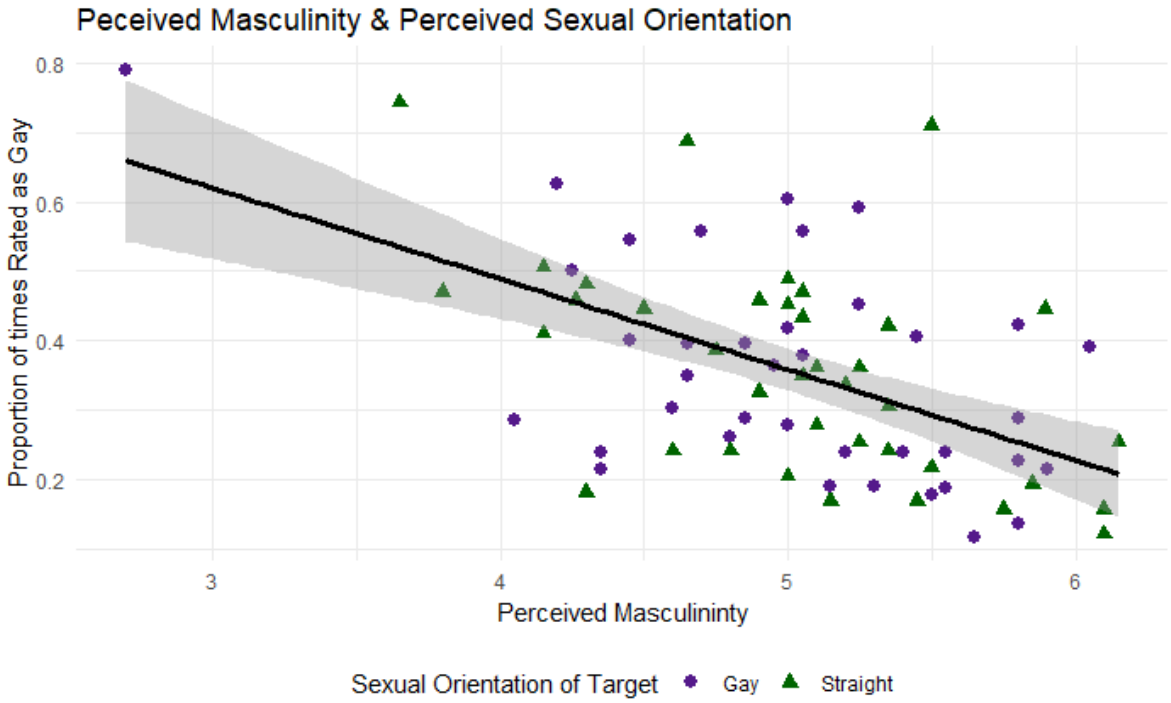


*Figure 5.* Perceived masculinity and perceived sexual orientation

## Discussion

Overall, people were not able to identify the *gay* faces at above chance accuracy. The main finding of the present study is that perceived masculinity/femininity, but not actual sexual orientation predicts perceived sexual orientation. These finding suggest that people do not have accurate gaydar abilities, despite the perception that they are accurate. Although overall accuracy, independent of condition (experimental/real-world) and face type (gay/straight), was significantly above chance, participants still only performed around 53% accuracy, which may not represent a meaningful difference.

Previous research (e.g. Rule & Ambady, 2008; Rule et al. 2008) has shown above chance accuracy for identifying gay faces, which this study does not replicate. In the current study, participants were unable to accurately identify the sexual orientation of gay men in either condition (experimental or real-world). Participants who saw only three gay faces (i.e., those in the real-world condition) tended to have higher accuracy for gay faces than did those who saw 36 gay faces (i.e., those in the traditional experimental condition), but neither group performed significantly better than chance. The primary difference between the current study and previous literature is the utilization of standardized images. Standardized images remove contextual cues (e.g. expression) and allow for judgements based on face shape alone. If gaydar is people accurately and reliably picking up on legitimate differences in face shape, the standardization of images and the distribution of faces should not have an effect on accuracy. Therefore these results suggest that gaydar judgements are not accurate.

Unlike Brambilla et al. (2012), contact was not found to affect accuracy of categorization. Brambilla et al. found that those with more contact were less confident but more accurate, however, the current study did not find a relationship between accuracy and contact or confidence, and/or their interaction. This finding casts doubt on the relationship between contact and gaydar accuracy. The argument could be made that this is due to a lack of contact with the Czech faces, given that I used Northern Californian raters and Czech targets. However, gaydar has previously been suggested to be accurate across cultures, even when directly comparing American and Czech men (Valentova et al. 2011), so this explanation is unlikely to resolve the discrepancies between the current findings and those of previous research.

The finding that more feminine faces were rated as gay more often in the current study is in line with Freeman et al. (2010) as well as Valentova et al. (2014). It seems that the stereotype that gay men are feminine may be the driving factor in people's sexual orientation judgements, suggesting that so-called gaydar really reflects the perception of sex typicality in faces rather than actual morphological differences in the faces of gay and straight men. However, as the gay men's faces were not all rated as extremely feminine and the straight men's faces were not all rated as extremely masculine (see Figure 1 above), sex typicality may not always be a helpful heuristic. For example, the face that was rated as gay the most often (79% of the times seen), upon visual inspection, presented both feminine and masculine features (e.g., rounder & feminine jaw but a strong masculine brow; Perrett et al. 1988), suggesting that using separate scales for both perceived masculinity and femininity instead might provide a more nuanced

understanding of the relationship between sexual dimorphism and perceptions of sexual orientation.

To visually inspect the morphometric differences between gay and straight faces, I also created a composite straight and composite gay face by averaging the 2D shape, color, and texture of the 38 gay men's faces and the 38 straight men's faces (Tiddeman, Burt, & Perrett, 2001) and the 10 faces rated as gay the most often and the 10 rated the gay least often were used to create composites of what raters perceived a gay and straight person should look like. The two faces made from using actual sexual orientation looked similar and it took strong transformations of a separate face to observe visible morphological differences between the composites (See Figure 6). It could be that participants were unable to correctly identify the gay individuals in the photo set because the gay men did not generally differ in appearance from the straight men. Whether the lack of visual facial differences is due to gay men and straight men not having reliable perceivable facial differences or the similarity is simply an artifact of the photo set is unclear. It should be noted that photos from Valentova and Havlíček (2013) were used as a subset of the images in the current study; in the previous (2013) study, the gay men's faces were rated as more masculine than the straight men's faces by a set of Czech raters. Further research examining gaydar accuracy using a different set of standardized images is necessary to tease out an answer.

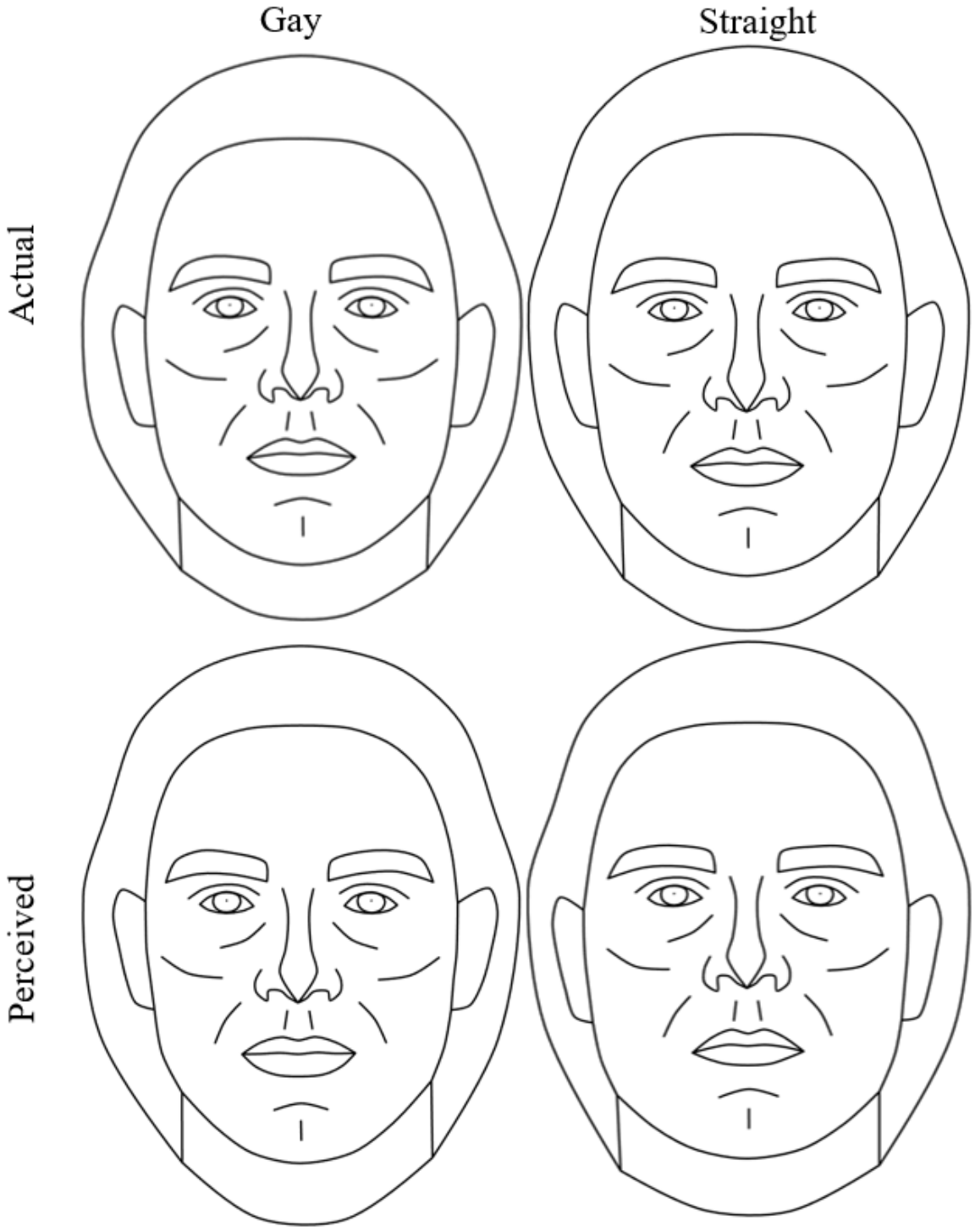


*Figure 6.* Averages of the actually and perceptually gay and straight faces

## Limitations

This study has several limitations. First, although every effort was made to use highly standardized images, it is difficult to access a large set of faces with corresponding sexual orientation information, so the number of faces I was able to use was limited. It should also be noted that some of the faces had some facial hair; this is not ideal for standardization and may have affected masculinity and sexual orientation perception of some of the faces. Additionally, due to the experimental setup on Qualtrics, to ensure that a random subset of faces was selected for the real-world condition, the two conditions saw an unequal number of total faces. However, again, if gaydar is something that people can do reliably and accurately, the total number of faces should not matter. Future research should focus on the collection of other standardized images to compare these results to.

## Conclusions

Past literature has suggested that people can accurately identify someone's sexual orientation, giving credibility to the folk notion of gaydar (Ambady, Hallahan & Conner, 1999; Rule et al 2008). This previous research has suggested that there may be phenotypic differences between straight and gay men. The current study used standardized images to investigate if sexual orientation could be determined from the face alone, isolating these potential morphometric differences, while endeavoring to remove as many additional contextual cues as possible. Results showed that people were not able

to identify the sexual orientation of the targets, even when the distribution of gay and straight faces more closely aligns with the population, calling into question the claim that there may be underlying morphological differences between gay and straight men's faces. The finding that more masculine men are rated as gay less often is replicated in the current study (Freeman et al 2010; Valentova et al. 2014). Results from this study suggest that people's sexual orientation judgments are based off of sex typicality of the face rather than actual morphological variation.

## References

Ambady, N., Hallahan, M. & Conner, B. (1999). Accuracy of judgements of sexual orientation from thin slices of behavior. *Journal of Personality and Social Psychology*, *77*, 538-547. doi: 10.1037/0022-3514.77.3.538

Berry, D. S. (1990). Taking people at face value: Evidence for the kernel of truth hypothesis. *Social Cognition, 8*, 343-361. doi:10.1521/soco.1990.8.4.343

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

Brewer, G., & Lyons, M. (2016). Discrimination of sexual orientation: Accuracy and confidence. *Personality and Individual Differences, 90*, 260-264, doi:10.1016/j/paid.2015.11.019

Cox, W. T. L., & Devine, P. G. (2014). Stereotyping to infer group membership creates plausible deniability for prejudice based aggression. *Psychological Science, 25*, 340-348. doi:10.1177/0956797613501171

Cox, W. T. L., Devine, P. G., Bischmann, A. A., & Hyde, J.S. (2016). Inferences about sexual orientation: The roles of stereotypes, faces, and the gaydar myth. *The Journal of Sex Research, 53*, 157-171. doi:10.1080/00224499.2015.1015714

Drydakis, N. (2015). Sexual orientation discrimination in the United Kingdom's labour market: A field experiment. *Human Relations, 68*, 1769 - 1796. doi: 10.1177/0018726715569855

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128,* 203 - 235 doi:10.1037/0033-2909.128.2.203

Eliášová, H., & Krsek, P. (2007) Superimposition and projective transformation of 3D object. *Forensic Science International, 167,* 146-153. doi:10.1016/j.forsciint.2006.06.062

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

Fruhen, L., Watkins, C. D., & Jones, B.C. (2015). Perceptions of facial attractiveness, dominance and trustworthiness predict managerial pay awarded in experimental tasks. *Leadership Quarterly, 26,* 1005-1016. doi:10.1016/j.leaqua.2015.07.001

Gates, G. J. (2017, Jan 11). *In US, more adults identifying as LGBT*. Retrieved from http://news.gallup.com/poll/201731/lgbt-identification-rises.aspx

Goldstein, A. G. & Chance, J. E. (1985). Effects of training on Japanese face recognition: Reduction of the other race effect. *Bulletin of the Psychonomic Society, 23,* 211-214. doi:10.3758/BF03329829

Hahn, A. C., Holzleitner, I. J., Lee, A. J., Kandrik, M., O'Shea, K. J., DeBruine, L. M., & Jones, B. C. (2018) Facial masculinity is only weakly correlated with handgrip strength in young adult women. *bioRxiv*, doi:10.1101/425017

Hehman, E., Leitner, J. B., & Gaertner, S. L. (2013) Enhancing statics facial features increases intimidation. *Journal of Experimental Social Psychology, 49,* 747-754. doi:10.1016/j.jesp.2013.02.015

Holzleitner I. J., Lee A., Hahn A. C., Kandrik M., Bovet J., Renoult J. P., Simmons D., Garrod O., DeBruine, L. M., & Jones B. C. (2018). Comparing theory-driven and data-driven attractiveness models using images of real women's faces. *PsyArXiv*. doi: 10.31234/osf.io/vhc5k

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., Gill, S., Reichman, V., & Tassinary, L. G. (2007). Swagger, sway, and sexuality: Judging sexual orientation from body motion and morphology. *Journal of Personality and Social Psychology, 93*, 321-334. doi:10.1037/0022-3514.93.3.321

Kinsey, A. C., Pomeroy, W. B., Martin, C. E., & Gebhard, P. H. (1953). Sexual behavior in the human female. Philadelphia: W. B. Saunders Company.

Knipe, E. (2017, Oct 4). *Sexual identity, UK:2016.* Retrieved from https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/sexuality/bulletins/sexualidentityuk/2016

DeBruine, L. (2017, December 4). Webmorph (Version v0.0.0.9001). Zenodo. doi:10.5281/zenodo.1073696

LeVay, S. (1991) A difference in hypothalamic structure between heterosexual and homosexual men. *Science*, 253, 1034 - 1037. doi: 10.1126/science.1887219

Lindsay, D. S., Jack, P. C., & Christian, M. A. (1991). Other-race face perception. *Journal of applied psychology, 76*(4), 587. doi:10.1037/0021-9010.76.4.587

Lyons, M., Lynch, A., Brewer, G. & Bruno, D. (2014). Detection of sexual orientation ("gaydar") by homosexual and heterosexual women. *Archives of Sexual Behavior, 43,* 345-352, doi:10.1007/s10508-013-0144-7

Martin, J. T. & Nguyen, D. H. (2004). Anthropomorphic analysis of homosexuals and heterosexuals: Implications for early hormone exposure. *Hormones and Behavior,* 45, 31 - 39. doi:10.1016/j.yhbeh.2003.07.003

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7,* 3-35. doi:10.1037/1076-8971.7.1.3

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *PNAS, 105,* 11087-11092. doi:10.1073/pnas.0805664105

Penton-Voak, I. S., Pound, N., Little, A. C., & Perrett, D. I. (2006). Personality judgments from natural and composite facial images: More evidence for a "kernel of truth" in social perception. *Social cognition, 24,* 607-640. doi:10.1521/soco.2006.24.5.607

Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., ... & Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature, 394,* 884. doi:10.1038/29772

Plöderl, M. (2014). Bayesian advice for gaydar-based picking up: Commentary on Lyons, Lynch, Brewer, and Bruno. *Archives of Sexual Behavior, 43,* 7-9. doi:10.1007/s10508-013-0178-x

Rice, W. R., Friberg, U., & Gavrilets, S. (2012). Homosexuality as a consequence of epigenetically canalized sexual development. *The Quarterly Review of Biology, 87*, 343-368. doi:10.1086/668167

Rieger, G., Linsenmeier, J. A. W., Gygax, L., Garcia, 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

Robertson, J. M., Kingsley, B. E., & Ford, G. C. (2019). Psychometric and faciometric support for observable facial feminization in gay men. *Journal of Homosexuality, 66*, 349 -367. doi: 10.1080/00918369.2017.1411692

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

Rule, N. O., Ambady, N., Adams, R. B., Jr., & Macrae, C. N. (2008). Accuracy and awareness in 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. (2009). 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*, 1499-1507. doi:10.1177/0146167211415630

Rule, N. O., Johnson, K. L., & Freeman, J. B. (2016). Evidence for the absence of stimulus quality differences in tests of the accuracy of sexual orientation judgements: A reply to Cox, Devine, Bischmann, & Hyde (2016). *The Journal of Sex Research*. *54,* 813-819. doi:10.1080/00224499.2016.1205547

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*, 881-886. doi:10.1177/0956797611412394

Rule, N. O., Tskhay, K. O., Brambilla, M., Riva, P., Andrzejewski, S. A., & Krendl, A. C. (2015). The relationship between anti-gay prejudice and the categorization of sexual orientation. *Personality and Individual Differences*, *77*, 74-80. doi:10.1016/j.paid.2014.12.038

Savins-Williams, R. C. (2006) Who's gay? Does it matter? *Current Directions of Psychological Science*, *15*(1), 40-44. doi:10.1111/j.0963-7214.2006.00403.x

Skorska, M. N., & Bogaert, A. F. (2017). Sexual orientation, objective height, and self-reported height. *The Journal of Sex Research*, 54, 19-32, doi: 10.1080/00224499.2015.1124831

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

Stephen, I. D., Smith, M. J. L., Stirrat, M. R., & Perrett, D. I. (2009). Facial skin coloration affects perceived health of human faces. *International Journal of Primatology*, *30*, 845-857. doi:10.1007/s10764-009-9380-z

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

Tiddeman, B., Burt, M., & Perrett, D. (2001). Prototyping and transforming facial textures for perception research. *IEEE computer graphics and applications, 21*, 42-50. doi:10.1109/38.946630

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in cognitive sciences, 12*, 455-460. doi:10.1016/j.tics.2008.10.001

Toma, C. L., & Hancock, J. T. (2010). Looks and Lies: The Role of Physical Attractiveness in Online Dating Self-Presentation and Deception. *Communication Research, 37*, 335–351. https://doi.org/10.1177/0093650209356437

Třebický, V., Fialová, J., Kleisner, K., & Havlíček, J. (2016) Focal length affects depicted shape and perception of facial images. *PLoS ONE, 11*, 1-14. doi:10.1371/journal.pone.0149313

Valentova, J. V., & Havlíček, J. (2013) Perceived sexual orientation based on vocal and facial stimuli is linked to self-rated sexual orientation in czech men. *PLoS One, 8*, 1-6. doi:10.1371/journal.pone.0082417

Valentova, J. V., Kleisner, K., Havlíček, 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., Havlíček, 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/s1508-011-9818-1

Walker, P. M., & Tanaka, J. W. (2003). An encoding advantage for own-race versus other-race faces. *Perception, 32*, 1117-1125. doi:10.1068/p5098

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, 246-257. doi:10.1037/pspa0000098

Witty, M. T. (2008). Revealing the 'real' me, searching for the 'actual' you: Presentations of the self on an internet dating site. *Computers in Human Behavior, 24*, 1707 - 1723. doi:10.1016/j.chb.2007.07.002