HHS Public Access

Author manuscript

J Sex Res. Author manuscript; available in PMC 2016 February 01.

Published in final edited form as:

J Sex Res. 2016 February ; 53(2): 157–171. doi:10.1080/00224499.2015.1015714.

Inferences About Sexual Orientation: The Role of Stereotypes, Faces, and The Gaydar Myth

William T. L. Cox, **Patricia G. Devine**, **Alyssa A. Bischmann**, and **Janet S. Hyde**

Psychology Department, University of Wisconsin – Madison

Abstract

In the present work, we investigate the pop cultural idea that people have a sixth sense, called "gaydar," to detect who is gay. We propose that "gaydar" is an alternate label for using stereotypes to infer orientation (e.g., inferring that fashionable men are gay). Another account, however, argues that people possess a facial perception process that enables them to identify sexual orientation from facial structure (Rule et al., 2008). We report five experiments testing these accounts. Participants made gay-or-straight judgments about fictional targets that were constructed using experimentally-manipulated stereotypic cues and real gay/straight people's face cues. These studies revealed that orientation is not visible from the face—purportedly "face-based" gaydar arises from a third-variable confound. People do, however, readily infer orientation from stereotypic attributes (e.g., fashion, career). Furthermore, the folk concept of gaydar serves as a legitimizing myth: Compared to a control group, people stereotyped more when led to believe in gaydar, whereas people stereotyped less when told gaydar is an alternate label for stereotyping. Discussion focuses on the implications of the gaydar myth and why, contrary to some prior claims, stereotyping is highly unlikely to result in accurate judgments about orientation.

Keywords

orientation; stereotypes/stereotyping; gaydar; antigay prejudice; feminist perspectives

Inferences About Sexual Orientation: The Role of Stereotypes, Faces, and The Gaydar Myth

> "I am not gay. I am, however, thin, single, and neat. Sometimes when someone is thin, single, and neat, people assume they are gay because that is a stereotype. They normally don't think of gay people as fat, sloppy, and married."
>
> — Jerry Seinfeld (1993)

Seinfeld's comedic commentary on the use of social stereotypes to infer orientation is perceptive. The idea that people use stereotypic cues to make inferences about group membership has a long history in the stereotyping literature (e.g., Allport, 1954; Cox &

Address correspondence to William Cox, Department of Psychology, University of Wisconsin – Madison, 1202 W Johnson St, Madison, WI, 53706. will.cox@me.com Phone: (353) 275-1223.

The data and experimental materials from this article are available publicly at http://www.sciencecox.com/pub/tgm15

Devine, 2014; Dotsch, Wigboldus, & van Knippenberg, 2011), but Seinfeld's observation also reflects a broader cultural notion that people believe they have a special ability to determine who is gay or lesbian. In popular culture, this special ability has been referred to as "gaydar," a kind of sixth sense or intuition that uniquely informs orientation judgments (Bronski, Pellegrini, & Amico, 2013; Reuter, 2002; Shelp, 2002). People's self-reported experience with gaydar (Shelp, 2002) and the pop cultural framework of gaydar (e.g., Reuter, 2002) indicate that gaydar seems to rely primarily on stereotypic attributes, like fashion or career choices.

A *stereotype* is a learned association between two social concepts that do not define each other (e.g., *fashionable* → *gay*) (Cox, Abramson, Devine, & Hollon, 2012). Stereotypes are activated automatically, and people are not always fully aware of how stereotypes inform their inferences (Devine, 1989). Similar to the notion of "gaydar" as a special intuition for detecting gay people, stereotypes often operate under the guise of an intuition that aids quick judgments about others (Pettigrew, 1979; Uhlmann & Cohen, 2007). For these reasons, we propose that the folk concept of "gaydar" may serve as a legitimizing myth. Legitimizing myths are widespread, seemingly innocuous beliefs or ideologies that grant apparent legitimacy to processes like stereotyping, prejudice, and oppression (Pratto, Sidanius, Stallworth, & Malle, 1994). Specifically, in the present work we tested whether *the gaydar myth* perpetuates the use of gay stereotypes by giving the stereotyping process a more socially and personally acceptable label.

It is important to note that the term "myth" itself does not indicate truth or falsity; it merely suggests that the idea—in this case, the idea that people have gaydar—is widely known and believed by many as a self-apparent truth (Pratto et al., 1994). And, in fact, there may be some truth to the gaydar myth: Some recent work suggests that people can accurately identify orientation from facial structure (e.g., Rule, Ambady, Adams, & Macrae, 2008), and others have argued that stereotypes yield accurate conclusions about orientation (e.g., Rieger, Linsenmeier, Gygax, Garcia, & Bailey, 2010). We will call these inference processes *face-based gaydar* and *stereotype-based gaydar*, respectively. In what follows, we review evidence relevant to both of these accounts and then report a series of experiments examining the extent to which face cues and stereotypic cues independently or jointly influence orientation inferences. We go on to test whether and how these inference processes are affected by manipulating belief in the gaydar myth. Lastly, we address issues of judgment accuracy, and lay out why neither stereotypes nor face cues are likely to give rise to pragmatically accurate inferences about orientation.

Faces and Stereotypes as Cues to Orientation

According to some recent work, gay and straight men's faces, and lesbian and straight women's faces, are distinct in visibly perceptible ways that can be used by perceivers to accurately identify the orientation of others (e.g., Rule & Ambady, 2008; Rule et al., 2008; Rule, Ambady, & Hallett, 2009). Although the effects are small, this face-based gaydar has been demonstrated several times and occurs even in situations with impoverished cues, such as when only small portions of the face are shown (e.g., just the eyes, nose, or mouth; Rule et al., 2008), and under extremely brief exposure time (i.e., when the face is flashed for only

50ms; Rule & Ambady, 2008). Rule and colleagues argue that this process is automatically activated (Rule, Macrae, & Ambady, 2009) and that it occurs across cultures (Rule, Ishii, Ambady, Rosen, & Hallett, 2011).

The alternative account of stereotype-based gaydar is more consistent with Seinfeld's observation and suggests that social stereotypes play a critical role in perceivers' orientation inferences. Indeed, we suggest that in addition to the classically-emphasized functions of stereotypes in culture, such as justifying prejudice and oppression (Devine & Sherman, 1992; Pratto et al., 1994; Riggs, 1987), stereotypes about gays and lesbians serve a categorization function (see also Russo, 1987). According to this account, because the defining features of gay/lesbian group membership (e.g., gay/lesbian identity, same-sex attraction) are not highly visible, gay and lesbian stereotypes developed—and persist—in culture specifically to create a set of highly visible cues that make it *seem* that these social groups are visibly identifiable. This function is readily apparent in the media, which has a history of using stereotypic cues to imply that a character is gay or lesbian (Cartei & Reby, 2012; Dennis, 2009; Russo, 1987). Indeed, an abundance of self-report, correlational, and experimental evidence has shown that people rely on stereotypic attributes, such as fashion, hairstyle, or femininity/masculinity, to make judgments about orientation (self-report: Matthews & Hill, 2011; Shelp, 2002; correlational: Ambady, Hallahan, & Conner, 1999; Freeman, Johnson, Ambady, & Rule, 2010, Studies 2 & 3; Gaudio, 1994; Johnson, Gill, Reichman, & Tassinary, 2007, Study 3; Rieger et al., 2010; Smyth, Jacobs, & Rogers, 2003; Van Borsel, & Van de Putte, 2014; experimental: Cox & Devine, 2014; Dotsch et al., 2011, Study 3; Freeman et al., 2010, Study 1; Johnson et al., 2007, Studies 1 & 2).

Some have speculated that the face-based gaydar discussed earlier may itself reflect a form of stereotyping (Freeman et al., 2010, p. 1328), to the extent that it depends on stereotypical information displayed on the face (e.g., facial grooming, emotional expressiveness). Because the stimulus pictures used in face-based gaydar research are retrieved from online dating websites, it is possible that such self-presentation could occur. As such, face-based gaydar and stereotype-based gaydar are not necessarily mutually exclusive inference processes. Rule and Ambady have argued, however, that face-based gaydar does not arise from stereotyping processes. These claims are rooted in evidence that gay and straight men in fact self-present *counter*-stereotypically on their dating profiles (i.e., gay men self-present as more masculine, straight men as more feminine; Rule & Ambady, 2008, p. 1103). In this case, face-based gaydar may reflect phenotypic face differences—similar to secondary sex characteristics that differentiate men and women's faces. To date, then, researchers have provided conflicting speculations about the mechanism underlying face-based gaydar. One goal of the present work was to address and resolve this apparent conflict.

The Present Work

Our first step was to determine what gives rise to the judgments attributed to "gaydar." In other words, what cues have the most influence on people's judgments about orientation? Study 1 served that end by assessing the relative influence of actual face cues and manipulated stereotypic cues on orientation judgments. Unexpectedly, Study 1 revealed a potential confound that may explain the mechanism of face-based gaydar. This confound

was subsequently tested for male targets in Studies 2 and 3, and extended to female targets in Study 4. Finally, in Study 5 we manipulated participants' beliefs about gaydar to explore whether the gaydar myth legitimizes stereotyping processes.

Studies 1A and 1B

In both Studies 1A and 1B, participants made gay-or-straight judgments about fictitious stimulus men. Each fictitious target was constructed using a descriptive statement and a picture of a real White man who identified either as gay or straight. Stereotype content was manipulated using the statements, which were gay-stereotypic, stereotype-neutral, or straight-stereotypic. Thus, our fictitious targets possessed both experimentally-manipulated stereotypic cues as well as whatever cues are present in pictures of real gay and straight men's faces. For each study, the design was 2 (Picture: Gay vs. Straight) X 3 (Statement: Gay-Stereotypic vs. Stereotype-Neutral vs. Straight-Stereotypic) within-subjects. Participants therefore judged a mix of six types of fictitious target men (men constructed with a gay man's picture and a gay-stereotypic statement, men constructed with a gay man's picture and a stereotype-neutral statement, etc.). To assess stereotyping in this paradigm, we examined participants' rates of stereotype-congruent judgments (i.e., how often they relied on the statements and guessed that men described with gay-stereotypic statements were gay and men described with straight-stereotypic statements were straight). The stereotype-neutral trials provided an opportunity to replicate previous work on face-based gaydar.

Most prior work on face-based gaydar has employed pictures of the faces with hairstyle. Hairstyle, however, is a highly controllable non-face cue that alters the perception of faces (see Brebner, Martin, & Macrae, 2009; DeBruine, Jones, Smith, & Little, 2010). Using stimuli without hairstyle provides a better opportunity to evaluate effects driven by phenotypic—rather than self-presentational—differences between gay and straight men's faces. Therefore, we used pictures with the face only in Study 1A and face with hairstyle in Study 1B.

Method

Participants—In Study 1A, 53 undergraduates (39 female, 47 White, all straight) participated online for extra credit. In Study 1B, 46 undergraduates (25 female, 43 White, 45 straight) participated online for extra credit.[1] In Studies 1, 2, 4, and 5, data were not collected from people who had been in the U.S. fewer than 4 years, to make reasonably sure people were acculturated to U.S. stereotypes of gay and straight men.

Pictures—Our original intent was to use the pictures from Rule and colleagues' (2008) research to allow for a precise replication of their findings. Our request for those stimuli was denied, however, citing concerns about protecting the identities of the target men (N. O. Rule, personal communication, March 1, 2011). We therefore collected our own set of pictures, closely following the retrieval strategy outlined by Rule et al. (2008). Specifically, we obtained public domain, freely available photographs of White men who self-reported

[1]The sample sizes reported in the present article excluded 13 participants who provided response sets, judging every target as gay or every target as straight, and 4 participants for whom the majority of their latencies were under 100ms, indicating task inattention.

being “gay” (n = 57) or “straight” (n = 52). We gathered pictures of men from both a Midwestern state and an Eastern state in the United States, each far from the state in which data were collected. All pictures came from a single dating website that serves people of any orientation. We selected pictures from the first profiles retrieved by searches in these locations. We did not collect pictures when the profile owner’s face was not clearly visible (i.e., fuzzy pictures or pictures without the person facing the camera) or in which the profile owner had facial piercings. We cropped the pictures, standardized their sizes, and placed them on a white background, as shown in Figure 1. Four pictures from the set obtained (2 gay and 2 straight) were excluded because the men’s hairstyles covered part of their faces, leaving us with 55 gay men’s pictures and 50 straight men’s pictures.

To norm the pictures, undergraduate raters (N = 27) blind to the orientation of the target men rated how angry, happy, sad, disgusted, fearful, neutral, and attractive each man was based on his face-and-hairstyle picture, from 1 (*not at all*) to 7 (*very*). These raters also judged the overall quality of each picture (1 = *very poor quality*, 7 = *excellent quality*). The only difference on these ratings was that the gay men’s pictures were higher quality than the straight men’s pictures, $t = 5.979$, $p < 0.001$, $d = 0.23$. These quality ratings had very high inter-rater reliability, $\alpha = 0.974$. Because retaining this difference in quality could have created unintended systematic variance in participants’ judgments, we selected 21 pairs of gay and straight pictures that were matched by their mean quality ratings. To be selected as a pair, the mean quality ratings of the gay picture and straight picture were at most 0.04 different from each other. After quality-matching in this way, there were no mean differences in the selected pictures’ ratings of quality, attractiveness, or emotional expressions, all $|t$’s$| \leq 1.229$, p’s ≥ 0.227.

Statements—To obtain gay-stereotypic, stereotype-neutral, and straight-stereotypic statements, we generated potential statements, relying on past literature (e.g., Kite & Deaux, 1987), the media (e.g., Reuter, 2002), and personal experiences, and we elicited stereotypic cues from undergraduates using discussion and free response questions (e.g., “How would you be able to tell if a man was gay without being able to ask him?”). This yielded a list of 130 potential statements. Thirty-six undergraduates (after dropping 1 participant who responded with the midpoint for every item) rated these statements in a random order on a 7-point Likert scale (1 = *very gay*, 4 = *neither gay nor straight*, 7 = *very straight*). We used these ratings to choose the statements, selecting statements that relayed information likely to be found on an online social networking profile, to fit our cover story (see all chosen statements in Table 1).

Procedure—The present studies were run using Inquisit (2003). Participants read on the computer screen that they were participating in a study on interpersonal perception and that we “retrieved pictures of real Gay and Straight men and information about them from their online profiles. We want to pretest how people perceive these men before using them in another study.” Although the men in the task were fictitious, having been fabricated by combining real men’s pictures with randomly-assigned descriptive statements, we told participants that the men were real to decrease the likelihood of demand characteristics influencing their responses. To place both types of information on equal footing in the task,

the instructions encouraged people to consider both the face and the "personal information" (i.e., statements) when making their judgments. The instructions also emphasized that "about half" of the men were gay, to counteract any prior expectations that more men are straight than gay. Finally, participants were encouraged to use their "gut instincts" and reminded that their responses were confidential.

Participants pressed "G" or "S" on the keyboard to categorize men as gay or straight. Reminders of the response keys remained onscreen for the duration of the experiment. For each participant, the program constructed 42 fictitious stimulus men by randomly and orthogonally pairing one of the statements with each picture. Each person saw every statement and every picture exactly once and saw exactly 7 trials of each of the 6 picture-type/statement-type combinations (e.g., every participant saw exactly 7 targets constructed using a gay man's picture and a gay-stereotypic statement, 7 with a gay man's picture and a stereotype-neutral statement, and so on). The program displayed the picture and the statement on the screen simultaneously, and the stimuli remained on the screen until the participant made a gay or straight categorization. Participants were given as much time as they wanted to make each judgment.

Categorization rates—We computed each participant's percentages of "gay" and "straight" judgments in each of the six statement-picture combinations. In Studies 1–4, it was arbitrary whether analyses used the gay or straight categorization rates, because each could be derived from the other (e.g., a 40% gay categorization rate corresponded to a 60% straight categorization rate). Unless otherwise specified, statistical analyses used the gay categorization rates.

Previous work on face-based gaydar has sometimes been analyzed using signal detection (e.g., Rule & Ambady, 2008; Rule et al., 2011). To enable full comparison with prior work, we report supplemental signal detection analyses in the Appendix. These analyses replicated the effects reported in text via ANOVA. Also, the patterns reported below did not change when we analyzed these data using mixed models that treated the picture/statement stimuli as random effects.

Results and Discussion

All means and standard deviations from Studies 1–5 are reported in Table 2. The data and experimental materials from this article are available publicly at www.sciencecox.com/pub/tgm15.

Study 1A—We conducted a 2 (Picture: Gay vs. Straight) X 3 (Statement: Gay-Stereotypic vs. Stereotype-Neutral vs. Straight-Stereotypic) repeated-measures ANOVA on the gay categorization rates. Contrary to the findings of previous research (e.g., Rule et al., 2008), participants judged targets constructed with gay men's pictures as gay no more often than they guessed gay for the targets constructed using straight men's pictures, F (1, 52) = 0.114, p = 0.737, $\eta^2 < 0.001$. The statements, however, did affect categorization rates, F (2, 104) = 51.612, $p < 0.001$, $\eta^2 = 0.403$, with people guessing "gay" much more often on gay-stereotypic trials than either stereotype-neutral or straight-stereotypic trials. There was no interaction between Picture and Statement, F (2, 104) = 0.034, p = 0.967, $\eta^2 < 0.001$.

Study 1B—Inclusion of hairstyle did not alter the pattern of results; Study 1B replicated Study 1A in every way. There was no main effect of Picture F (1, 45) = 1.221, p = 0.275, η^2 = 0.002, but there was a main effect of Statement, F (2, 90) = 33.454, $p < 0.001$, η^2 = 0.329, and there was no Picture X Statement interaction, F (2, 90) = 0.323, p = 0.725, η^2 = 0.001.

In Studies 1A and 1B, people relied on stereotypic cues to draw conclusions about orientation, whether the fictitious targets were constructed with the picture of a gay or straight man. Contrary to expectations based on prior work, face cues did not influence categorization rates. Because this pattern is inconsistent with Rule and colleagues' (2008) work, we made efforts to identify the source of the disparity between their findings and ours. Aside from pairing the pictures with statements (which will be addressed by Study 3), a close examination of both sets of studies revealed one potentially important difference between our stimulus selection method and the method used by Rule et al. (2008). That is, although we duplicated their stimulus *retrieval* procedure, we subsequently matched the gay and straight men's pictures for quality. No prior work on face-based gaydar has ever used quality-matched pictures of real men, nor even reported testing the quality of the retrieved pictures (e.g., Freeman et al., 2010, Studies 2 & 3; Lyons, Lynch, Brewer, & Bruno, 2014; Rule & Ambady, 2008; Rule et al., 2008; Tabak & Zayas, 2012). In the following analyses, we examined whether picture quality relates to participants' categorization of the targets.

Picture quality analyses—To examine the effect of picture quality on categorization rates, we computed gay categorization rates by *picture*, (rather than by *participant*, as in the previous analyses). Each picture's categorization rate, therefore, reflects how often participants concluded that target men constructed using that picture were gay (separating these analyses by statement type made no difference). The pictures' gay categorization rates and mean quality ratings positively correlated ($r_{Study1A}$ = 0.284, p = 0.069; $r_{Study1B}$ = 0.260, p = 0.096; N = 42), suggesting that people were more likely to guess that higher-quality pictures were of gay men.

Picture quality, therefore, may be a source of systematic variance that could create effects like those observed by Rule and colleagues. To explore this possibility, we conducted Studies 2A (faces) and 2B (faces and hairstyles) using pictures that retained the picture quality difference we found between gay and straight men's dating profile pictures online.

Studies 2A and 2B

Method

Participants—In Study 2A, participants were either psychology undergraduates who participated in person (n = 183) or were unpaid volunteers recruited via Craigslist.com who participated online (n = 68), yielding 251 participants (143 female, 194 White, 205 straight). Study 2B had 97 participants in total (64 female, 79 White, 83 straight, 27 recruited online). There were no differences between the student and non-student samples in either study | M_{diff}'s| ≤ 0.006, |t's| ≤ 0.334, p 's ≥ 0.620. White, female experimenters conducted the experimental sessions in these and the following studies that involved in-lab participation.

Pictures—In Study 2 we used the pictures rated highest in quality within each category (21 gay, 21 straight). Matching the pattern in the full set of pictures, the gay men's pictures were higher in quality than the straight men's, $t = 9.941$, $p < 0.001$, $d = 0.60$.

Results

Study 2A—As in Study 1, we conducted a 2 (Picture) X 3 (Statement) repeated-measures ANOVA. Replicating past work, people guessed "gay" more for targets constructed using gay men's pictures than straight men's pictures, F (1, 250) = 144.997, $p < 0.001$, $\eta^2 = 0.042$. Replicating Study 1, there was also a main effect of Statement, F (2, 500) = 157.851, $p < 0.001$, $\eta^2 = 0.283$. There was no interaction between Picture and Statement, F (2, 500) = 1.589, $p = 0.205$, $\eta^2 = 0.001$, indicating that the main effects were largely independent.

Study 2B—Study 2B, which used the face-with-hairstyle pictures, replicated Study 2A in every way. There were main effects of Picture F (1, 96) = 77.281, $p < 0.001$, $\eta^2 = 0.071$, and Statement, F (2, 192) = 51.851, $p < 0.001$, $\eta^2 = 0.238$, but no interaction, F (2, 192) = 0.447, $p = 0.640$, $\eta^2 = 0.001$. Adding hairstyle to the picture stimuli made no difference.

Picture quality mediation analyses—The present analyses used three variables: the orientation of the pictured man (straight = 0, gay = 1), each picture's mean gay categorization rate, and each picture's mean quality rating. The pictures' gay categorization rates correlated with orientation ($r_{Study2A} = 0.416$, $p = 0.006$; $r_{Study2B} = 0.411$, $p = 0.007$), as would be expected by our ANOVA analyses and prior work. The quality ratings correlated with both orientation ($r = 0.545$, $p < 0.001$) and the gay categorization rates ($r_{Study2A} = 0.470$, $p = 0.002$; $r_{Study2B} = 0.442$, $p = 0.003$). Linear regressions predicting the categorization rates revealed that when controlling for the effect of picture quality ($\beta_{Study2A} = 0.347$, $p = 0.042$; $\beta_{Study2B} = 0.309$, $p = 0.071$), the effect of the pictured man's orientation dropped to nonsignificance ($\beta_{Study2A} = 0.227$, $p = 0.175$; $\beta_{Study2B} = 0.243$, $p = 0.153$).[2]

Discussion

Replicating Study 1, the statements accounted for 24–28% of the variance in orientation inferences. The effect of picture, however, accounted for only 4–7% of variance. These data suggest that people rely predominantly on gay-stereotypic information, rather than face cues, to conclude that men are gay. We resume our assessment of stereotype-based gaydar in Study 5, after a short detour to address the picture quality confound.

Although Study 2 replicated prior demonstrations of face-based gaydar, our internal analyses revealed that this effect was driven by picture quality differences, not facial differences between gay and straight men. Studies 1 and 2, however, are not directly comparable to prior work due to our inclusion of statements. It is possible that the presence

[2]Prior work (Freeman et al., 2010) has claimed that the face-based gaydar effect arises due to gay men's faces being more feminine than straight men's. Because facial "femininity" may affect quality ratings or vice versa, we asked 20 undergraduates to rate the femininity of the faces on a 7-point scale (1 = *not at all feminine*, 7 = *very feminine*), then conducted analyses using the mean femininity rating of each picture. Femininity correlated with Orientation ($r = 0.294$, $p = 0.059$) and Quality ($r = 0.361$, $p = 0.019$). In a regression with both Quality and Femininity predicting Orientation, only Quality ($\beta = 0.504$, $p = 0.001$) predicted Orientation, not Femininity ($\beta = 0.112$, $p = 0.438$). This suggests that gay and straight men differ in their picture quality, which affects ratings of femininity, rather than their faces differing in femininity.

of any other cues, even stereotype-neutral cues, may dilute or weaken the pictures' influence on categorization (e.g., Nisbett, Zukier, & Lemley, 1981). To confirm that picture quality was the underlying mechanism, Study 3 evaluated these effects without manipulated stereotypic information.

Study 3

Method

We recruited participants via Amazon's MTurk, from both the U.S. (N = 49; 31 female, 37 White, 45 straight) and outside the U.S. (N = 60; 20 female; 30 East Asian, 17 South Asian, 11 White; 45 straight, 7 not reported). There were no differences between the U.S. and non-U.S. samples, $|t$'s$| \leq 1.58$, p 's ≥ 0.117. Participants categorized gay and straight men's face pictures without stereotypic statements. We randomly assigned participants to one of two conditions: They judged either the quality-matched pictures from Study 1A or the quality-unmatched pictures from Study 2A.

Results and Discussion

We conducted a 2 (Quality: Matched vs. Unmatched) X 2 (Picture: Gay vs. Straight) mixed ANOVA. There was no main effect of Quality, F (1, 107) = 1.152, p = 0.220, η^2 = 0.014. The main effect of Picture, F (1, 107) = 5.505, p = 0.021, η^2 = 0.047 was qualified by the predicted Quality X Picture interaction F (1, 107) = 4.800, p = 0.031, η^2 = 0.041. See Figure 2, Top. Only the quality-unmatched pictures produced higher gay categorization rates for gay than straight targets, t (61) = −3.358, $p < 0.001$, d = 0.36, replicating Study 2A and prior work using stimuli that were not matched on quality. Replicating Study 1A, there was no difference in gay categorization rates for the quality-matched gay and straight pictures, t (46) = −0.107, p = 0.915, d = 0.01. Even without stereotypic statements, the "face-based" gaydar effect only arose when the stimuli had not been matched on quality.

It appears that "face-based" gaydar arises as an artifact of the quality of the dating website pictures and that orientation is not visible from the face, for male targets at least. To complete our assessment of purportedly "face-based" gaydar, in Study 4 we tested whether a picture quality confound might drive this effect for female targets as well.

Study 4

Stimuli

We collected pictures of self-identified straight (n = 70) and lesbian women (n = 70) and cropped them, following the procedure reported in Study 1 (see also Rule et al., 2009). We excluded 13 straight targets and 9 lesbian targets because their faces were not fully visible due to their hairstyles or how they had initially cropped their pictures. This left 61 lesbian women's pictures and 57 straight women's pictures for a total of 118 pictures.

As in Study 1, undergraduates' (N = 45) ratings of the photos revealed a quality difference, with the lesbians' pictures being of higher quality on average than the straight women's, t (116) = 2.228, p = 0.029. We selected 39 pairs of quality-matched pictures, and within this

matched set, the lesbian and straight women's pictures did not differ on quality t (76) = −0.282, p = 0.779, or any of the other rated variables, all |t's| ≤ 1.397, p's ≥ 0.166.

Participants and Procedure

Ninety-six undergraduates (66 female, 75 White, 94 straight) participated online for extra credit. Participants were randomly assigned to make lesbian-or-straight judgments about the set of quality-matched pictures or the full set of quality-unmatched pictures. Participants completed a modified version of Study 3's task, with the phrases "gay men" and "straight men" replaced with "lesbian women" and "straight women" throughout. Participants pressed "L" or "S" to categorize the women as lesbian or straight, respectively.

Results and Discussion

We conducted a 2 (Quality: Matched vs. Unmatched) X 2 (Picture: Lesbian vs. Straight) mixed ANOVA. There was no main effect of Quality, F (1,93) = 2.342, p = 0.129, η^2 = 0.025, and the main effect of Picture, F (1, 93) = 40.591, p < 0.001, η^2 = 0.270, was qualified by the predicted Quality X Picture interaction, F (1, 93) = 16.995, p < 0.001, η^2 = 0.113. See Figure 2, Bottom. In the quality-unmatched condition, participants judged the lesbian women's pictures as lesbian more often than straight women's pictures, t (45) = −7.734, p < 0.001, d = 0.51, matching prior work using quality-unmatched stimuli (e.g., Rule et al., 2009). In the quality-matched condition, however, there was no difference in lesbian categorization rates for the lesbian versus straight women's pictures, t (48) = −1.540, p = 0.130, d = 0.11.

Extending the findings of Studies 1–3, the lesbian women's pictures on average were of higher quality than the straight women's, and participants could not differentiate between lesbian and straight pictures when they were matched on quality.

Picture-based gaydar—Our findings suggest that orientation is not visible from the face. We cannot definitively state whether picture quality is responsible for the effects reported in previous work that did not test or match quality, because we were unable to use Rule and colleagues' stimuli. Our findings, however, raise the possibility that the previously reported gaydar effect arises from differences in self-selected dating profile *pictures*, rather than any differences in *faces*. We find it difficult to generate a hypothesis to explain why quality matching could eliminate any real differences between the faces of gay men/lesbians and straight men/women (see also Todorov & Porter, 2014). Because we retrieved our stimuli following the same procedure as previous studies on picture-based gaydar, the current studies should be equally generalizable to real people's faces and pictures. The present studies converge upon the conclusion that people cannot identify orientation from the face.

We can only speculate about why these quality differences are present, although this pattern is plausible given other research on online dating. We know, for example, that members of different groups adopt different self-presentational goals on their online dating profiles, including picture selection (Hancock & Toma, 2009; Toma, Hancock, & Ellison, 2008; Rule & Ambady, 2008; cf. Gudelunas, 2012). Based on consideration of the male targets alone, some might jump to a gender inversion conclusion that points to gay men as more feminine

and therefore more concerned about their appearances than straight men. This interpretation runs counter to past evidence that gay men self-present as more masculine than straight men in their online profiles (Rule & Ambady, 2008, p. 1103). Given that female targets show this pattern as well, we favor instead a minority status explanation—because lesbian women and gay men are small minority groups, their dating pools are much smaller than those of their straight counterparts. Gay men and lesbian women must therefore be more competitive in their smaller dating pool, and selecting higher-quality pictures likely makes them more competitive. This explanation is also consistent with prior research that demonstrates picture-based gaydar across different cultures, given that these minority pressures should exist in any culture (e.g., Rule et al., 2011). Unpacking the reason for this picture quality difference was outside of the goals of the present article and will await future work. The focal conclusion to emphasize in the present context is that it appears social perceivers cannot make accurate inferences about who is gay/lesbian from the face. If true physiognomic face differences existed between straights and lesbians/gays, quality matching the stimuli would be unlikely to wash them out (see also Todorov & Porter, 2014).

The gaydar myth—We return now to our exploration of whether the folk concept of "gaydar" serves the role of a legitimizing myth (Pratto et al., 1994). Studies 1 and 2 provide clear evidence that people rely heavily on stereotyping to infer orientation. Because stereotypes are automatically activated and pop to mind unbidden (Devine, 1989), as noted earlier, they often operate under the guise of a compelling "intuition" (Pettigrew, 1979; Uhlmann & Cohen, 2007). We propose that, in the case of stereotyping to infer orientation, the term *gaydar* emerged as a special label for this "intuition." Giving these intuitions the more socially acceptable label of "gaydar" may obscure the fact that inferences about orientation arise from stereotypes, and thus lead to higher levels of stereotyping. If this is indeed the case, the cultural concept of gaydar serves the function of a legitimizing myth. We tested our specific proposal that the gaydar myth perpetuates stereotyping to infer orientation in Study 5.

Study 5

We manipulated participants' belief in gaydar before they completed the stereotyping categorization task from Studies 1B and 2B. We told one group of participants that gaydar is a real ability and told another group that gaydar is merely another term for stereotyping. These two groups were compared to a control group, yielding three between-subjects conditions (Gaydar Belief: "Gaydar is Real" vs. Control vs. "Gaydar is Stereotyping"). If the idea of gaydar serves the function of a legitimizing myth, then authenticating the gaydar myth by telling participants that "gaydar is real" should increase gay stereotyping relative to the control group, and dispelling the gaydar myth by telling participants that "gaydar is stereotyping" should decrease stereotyping relative to the control group. In other words, compared to the control participants, participants in the "Gaydar is Real" condition will be more likely, and participants in the "Gaydar is Stereotyping" condition will be less likely, to guess that fictional men described with gay-stereotypic statements are gay.

Method

Participants and design—In our laboratory, 233 undergraduate participants (162 female, 198 White, 225 straight) were randomly assigned to one of the three Gaydar Belief conditions (Gaydar is Real $n = 78$, Control $n = 76$, Gaydar is Stereotyping $n = 79$). They completed a modified version of the task from Studies 1B and Study 2B, judging fictitious men constructed with the statements and face-and-hairstyle pictures. We conducted Study 5 with both the quality-matched ($N = 135$) and quality-unmatched ($N = 98$) pictures to allow for another replication of the quality confound findings. This replication was successful, and these patterns were unaffected by the gaydar belief manipulation.[3] Because they were unrelated to the objective of Study 5, the picture effects are discussed no further.

After the manipulation, participants made judgments about one of three types of fictitious target men. The target men were either described with a gay-stereotypic statement (e.g., "He likes shopping."), a stereotype-neutral statement (e.g., "He likes to read."), or a straight-stereotypic statement (e.g., "He likes football."). The design of Study 5, therefore, was 3 (Gaydar Belief condition: Gaydar is Real vs. Control vs. Gaydar is Stereotyping) X 3 (Stimulus Men: Gay-Stereotypic vs. Stereotype-Neutral vs. Straight-Stereotypic).

Procedure—After introducing the study, in the "Gaydar is Real" condition, the experimenter said, "Studies like this have shown that 'gaydar,' the ability to sense whether a person is gay, is a real perceptual ability. People use many subtle cues to perceive that a man is gay—people really do have *gaydar!*" In the "Gaydar is Stereotyping" condition, the experimenter said, "Studies like this have shown that 'gaydar,' the ability to sense whether a person is gay, is not a real perceptual ability. People use stereotypes to assume that a man is gay – people don't have gaydar, they just *stereotype!*" In the Control condition, the experimenter said nothing about gaydar or past studies examining it.

After the manipulation, the procedure and experimental task followed Studies 1B and 2B, with the following exceptions. Unlike the prior studies, we did not claim we were pretesting new stimuli, because this would somewhat mismatch information given in the manipulation. We also did not ask participants to refrain from censoring themselves, because we wanted them to self-censor if they wished to do so (e.g., in the "Gaydar is Stereotyping" condition; cf. Devine, 1989). Also, we added a third response option to enable participants to refrain from guessing, if they were so inclined. This option allowed them to "Press the 'n' key if you have absolutely no idea whether the man is Gay or Straight." Reminders about all three response options were presented onscreen throughout the program. After the task, participants completed a manipulation check, "Do you believe that people have gaydar that helps them sense who is gay?" on a 7-point Likert scale (1 = *No, not at all*, 4 = *Not sure*, 7 = *Yes, absolutely*).

[3]We submitted the Study 5 gay categorization rates to a 2 (Quality: Matched vs. Unmatched) X 3 (Gaydar Belief condition: Gaydar is Real vs. Control vs. Gaydar is Stereotyping) X 2 (Picture: Gay vs. Straight) X 3 (Statement: Gay-Stereotypic vs. Stereotype-Neutral vs. Straight-Stereotypic) mixed ANOVA. The effect of Picture, F (1, 227) = 17.909, $p < 0.001$, was qualified by the Quality X Picture interaction, F (1, 227) = 18.497, $p < 0.001$, as in Studies 1–4. The main effects of Statement, F (2, 454) = 89.774, $p < 0.001$ and Gaydar Belief condition F (2, 227) = 2.510, $p = 0.083$ were qualified by the predicted Statement X Gaydar Belief condition interaction, F (4, 454) = 4.409, $p = 0.002$. There were no other main effects or interactions, F's ≤ 1.969, p's ≥ 0.162.

Results and Discussion

Manipulation check and "No Idea" responses—Responses to the belief in gaydar Likert-item varied by condition, F (2, 230) = 30.399, $p < 0.001$, $\eta^2 = 0.209$, matching the manipulation. Participants in the "Gaydar is Real" condition ($M = 5.01$, $sd = 1.294$) believed in gaydar more than Control participants ($M = 3.89$, $sd = 1.613$, $d = 0.77$) and participants in the "Gaydar is Stereotyping" condition ($M = 3.20$, $sd = 1.480$, $d = 1.30$). Participants in the "Gaydar is Stereotyping" condition believed in gaydar less than Control participants ($d = -0.45$). In each condition, we conducted one-sample t-tests comparing these ratings to the midpoint of the scale (4 = *not sure*). The ratings from participants in the "Gaydar is Real" condition were significantly higher than the midpoint, t (77) = 6.911, $p < 0.001$, $d = 0.78$, ratings from Control participants were not different from the midpoint, t (76) = −0.569, $p = 0.571$, $d = -0.07$, and ratings from participants in the "Gaydar is Stereotyping" condition were significantly lower than the midpoint, t (78) = −4.791, $p < 0.001$, $d = -1.87$.

People used the "No Idea" response option very infrequently ($M_{PercentUse} = 5\%$, $sd = 10.4$), and its use did not differ by manipulated Gaydar Belief condition, F (2, 230) = 0.043, $p = 0.958$, $\eta^2 < 0.001$. Inclusion of the "No Idea" response option did not change the overall patterns of gay categorization: Analyses in the Control condition perfectly replicated the patterns found in Studies 1B and 2B. This successful replication opposes the expectation that normative concerns might make participants unwilling to make categorization decisions and suggests that the patterns in Studies 1–4 did not arise as artifacts of the forced-choice nature of the categorization task. Even when not forced to choose, participants relied upon stereotypic cues to infer orientation.

Primary analyses—We submitted the gay categorization rates to a 3 (Gaydar Belief condition: Gaydar is Real vs. Control vs. Gaydar is Stereotyping) X 3 (Statement: Gay-Stereotypic vs. Stereotype-Neutral vs. Straight-Stereotypic) mixed ANOVA, with Gaydar Belief condition between-subjects, and Statement within-subjects. The main effects of Statement, F (2, 460) = 95.383, $p < 0.001$, $\eta^2 = 0.208$, and Gaydar Belief condition F (2, 230) = 2.513, $p = 0.083$, $\eta^2 = 0.006$, were qualified by the predicted Statement X Gaydar Belief condition interaction, F (4, 460) = 4.240, $p = 0.002$, $\eta^2 = 0.019$. To unpack this interaction, we conducted a one-way ANOVA on the gay categorization rates for each statement type, with Gaydar Belief condition as a 3-level factor. As predicted, the manipulation did not affect responses to the stereotype-neutral or straight-stereotypic statement trials, F's ≤ 1.557, p's ≥ 0.213, but did affect responses to the gay-stereotypic statement trials, F (2, 230) = 7.691, $p = 0.001$, $\eta^2 = 0.062$. Matching our hypotheses, people in the "Gaydar is Real" condition had higher gay categorization rates on the gay-stereotypic trials than those in either the Control condition t (152) = 2.322, $p = 0.022$, $d = 0.37$, or the "Gaydar is Stereotyping" condition t (155) = 3.834, $p < 0.001$, $d = 0.61$. Participants in the "Gaydar is Stereotyping" condition had slightly lower gay categorization rates on these trials than did control participants, t (153) = 1.585, $p = 0.113$, $d = -0.26$.

Another way to evaluate stereotyping in the present study is to examine the extent to which gay-stereotypic statements led more often to gay than straight judgments. Specifically, as shown in Figure 3, comparing the gay categorization rates and the straight categorization

rates on the gay-stereotypic statement trials further elucidates the effects of our manipulation. Replicating Studies 1 and 2, participants in the Control condition guessed that the men described by gay-statements were gay more often than straight, t (75) = 2.486, p = 0.015, d = 0.55. This difference was much larger in the "Gaydar is Real" condition, t (77) = 5.655, $p < 0.001$, d = 1.25, strongly supporting our hypothesis that the gaydar myth justifies and increases the use of gay-stereotypic cues to infer that men are gay. Promoting the idea of gaydar as a real perceptual ability led participants to rely more heavily on stereotyping. Dispelling the gaydar myth by identifying gaydar as a stereotyping process, however, undermined reliance on stereotypic cues to infer orientation: Participants in the "Gaydar is Stereotyping" condition were equally likely to guess that gay-stereotypic men were gay or straight, t (78) = 0.201, p = 0.841, d = 0.04. It is notable that simply identifying gaydar as a stereotyping process led to a reduction in stereotyping. This pattern is especially striking because the manipulation did not explicitly discourage stereotyping, and we offered an opportunity to "opt out" of a categorization decision (i.e., participants could use the "no idea" response, but did not). This pattern is consistent with models that implicate *awareness* of stereotypic influences as a key step to reducing bias (e.g., Devine, 1989; Devine, Forscher, Austin, & Cox, 2012).

General Discussion

The pattern of findings across five studies challenges scientific claims that people possess a face-based gaydar ability. In our studies, the ostensibly "face-based" gaydar effect was driven by picture quality differences that were confounded with orientation. Removing this quality confound, either by quality-matching (Studies 1, 3, 4, & 5) or statistically controlling for quality (Study 2), eliminated any relationship between the orientation of the pictured man or woman and the orientation inferences drawn by perceivers. Participants consistently relied on stereotyping to infer orientation (Studies 1, 2, and 5). Authenticating the gaydar myth led participants to stereotype more, and dispelling the gaydar myth by identifying gaydar as a stereotyping process led them to stereotype less (Study 5). The evidence provided in Study 5 indicates that the folk concept of gaydar serves as a legitimizing myth, promoting stereotyping to infer orientation by giving that stereotyping process the alternate label of "gaydar."

Inaccuracy of Stereotype-Based Gaydar

Stereotype-based gaydar may seem largely harmless, or even useful, to the extent that the stereotypes lead to mostly accurate conclusions. In fact, much prior work on stereotype-based gaydar (e.g., Johnson et al., 2007, Study 3; Rieger et al., 2010) has argued that stereotypes (e.g., of gay men being feminine) lead to accurate conclusions about orientation. We suggest, however, that this reasoning builds on invalid assumptions and that even a seemingly accurate stereotype will lead to mostly inaccurate conclusions (see also Chapman & Chapman, 1969; Judd & Park, 1993; Olivola, Funk, & Todorov, 2014).

Consider the statistic that gay men are three times more likely to be alcoholic than straight men (Substance Abuse and Health Services, 2013, p. 66). Knowing that strong relationship, would you assume that an alcoholic man was likely gay? We surmise that you would not,

because alcoholism is not a common stereotypic trait that is supported and perpetuated by the dominant cultural stereotypes of gay men. When stereotypes guide one's expectations, however, one falls prey to errors in reasoning. And, in the present case, the gaydar myth further facilitates these errors. Specifically, we suggest that prior work's claims about accurate stereotype-based gaydar commits two related logical fallacies (see also Plöderl, 2014). The first is a *fallacy of the converse*, in which people wrongly assume that a relationship in one direction (gay men like shopping) implies the converse relationship (men who like shopping are gay). The second is a *base rate fallacy*, in which people wrongly ignore base rate information (Kahneman & Tversky, 1973).

To illustrate, suppose the stereotype that gay men like shopping is highly accurate: 100% of gay men and only 10% of straight men enjoy shopping. If one then relies upon this stereotype to infer that men are gay, what are the odds the inference will be correct? Without correction for base rates, the apparent answer would be a 91% probability (i.e., 10:1 odds) of being correct, because gay men are ten times more likely to enjoy shopping than straight men. The reality, however, is that there is only a 34% probability (10:19 odds) of being correct. This is because the base rate of gay men in the population is extremely small, approximately 5% of the adult male population (Plöderl, 2014; Savin-Williams, 2006). Thus, even if gay men are ten times more likely to like shopping, men who like shopping are still twice as likely to be straight as to be gay, because straight men outnumber gay men 19 to 1 (see Figure 4).

Mathematically, in order for a stereotype, or any cue, to yield higher-than-chance accuracy about a target being gay, gay men must be over 20 times more likely to possess that characteristic than straight men. To date, no such large differences have ever been shown between gay and straight men, or lesbian and straight women. Work exploring stereotype accuracy or accurate stereotype-based gaydar has used roughly equal numbers of gay/lesbian and straight targets (e.g., participants make gay/straight judgments about a set of targets that are 50% gay men and 50% straight men). These methods may be apt for questions about what cues perceivers use to make categorization decisions (as in the present work), or for questions about similarities or differences between these social groups (but see Harry, 1986). These sorts of research designs cannot, however, inform our understanding of whether people can accurately identify who is gay or lesbian in daily life, unless the design has real-world base rates built into it (for extended versions of this argument, see Olivola & Todorov, 2010 and Plöderl, 2014). No work of which we are aware has demonstrated patterns that surpass the 20:1 ratio needed to yield pragmatic accuracy when adjusted for real-world base rates. From a purely mathematical standpoint, it is highly unlikely that stereotypes—or any other cues—could yield pragmatically accurate gaydar[4] (see also Hamilton & Gifford, 1976; Mitroff & Biggs, 2014).

[4]It is important to note here that gaydar is defined as the ability to *detect that someone is gay*. Although the term "gaydar" has sometimes been applied to the use of eye-gaze and other flirting behaviors (e.g., Nicholas, 2004), these behaviors involve detecting specific attraction or romantic interest, not one's general sexual orientation. If Man A initiates sustained eye contact with Man B, for example, Man B may correctly infer that Man A is attracted to him and therefore is gay or bisexual. Two men who fail to make eye contact, however, may be straight or may merely be unattracted to one another. Being able to correctly infer that someone is flirting or romantically interested in you is not the same as being able to detect someone's orientation regardless of his or her interest in you. We thank an anonymous reviewer for helping us to see the importance of explaining this distinction.

Further Implications of The Gaydar Myth and Stereotyping to Infer Orientation

The very notion of gaydar would require that gay men and lesbians each possess some common essence that differentiates them from their straight counterparts and makes their orientation perceptible. As such, "gaydar" reflects a perspective that assumes underlying intergroup differences. Although such intergroup differences assumptions are not always explicitly stated, they are common and often troublesome, corresponding to higher levels of prejudice and lower levels of concern for discrimination and between-group inequality (e.g., Haslam, Rothschild, & Ernst, 2002; Hegarty & Pratto, 2001; Jayaratne et al., 2006; Prentice & Miller, 2007). An *intergroup similarities perspective*, however, leads one to question why and how the socially constructed defining features of orientation status (i.e., attraction, identity) might create perceptible differences (see also Hyde's 2005 gender similarities hypothesis; DeLamater & Hyde, 1998). As has been argued by many previous theorists, any two social groups are likely much more similar than different, and any *actual* differences between two social groups should be much smaller than the *stereotyped* differences between those groups (see, e.g., Allport, 1954; APA, 2006; Devine & Sherman, 1992; Fiske, 2010; Herek, Kimmel, Amaro, & Melton, 1991; Hyde, 2005; Rothbart & Taylor, 1992; Steele, 1997).

In some ways, the idea of gaydar—being able to tell who is gay vs. straight—seems useful at best and harmless at worst. The very fact that it seems harmless, however, may ultimately be responsible for its most pernicious effects. Our findings from Study 5 demonstrated that belief in gaydar perpetuates the use of stereotypes to jump to conclusions about orientation. Left unchecked, these stereotypes can lead to a host of adverse consequences (e.g., Cox & Devine, 2014; Matthews & Hill, 2011; Nadal, 2013). For example, as has recently been demonstrated (Cox & Devine, 2014), using a stereotypic cue to privately infer group membership can free prejudice perpetrators from concerns about social condemnation for their prejudice, yielding higher levels of prejudice-based aggression. Thus, the gaydar myth may not only promote the use of stereotypes to make inferences about orientation, but may also indirectly promote discrimination—even aggression—based on these inferences.

More generally, stereotypes by definition involve narrow thinking, and they support oppressive processes like prejudice and discrimination (Cox et al., 2012; Devine, 1989; Fiske, 1998). Although the valence of many gay stereotypes can be considered positive (e.g., gay men are fashionable, creative), they still promote restrictive beliefs and constrain opportunities (see Cheryan & Bodenhausen, 2000; Czopp, 2008; Kay, Day, Zanna, & Nussbaum, 2013). In addition, we want to emphasize that although the *stereotypes*, as cultural or mental associations, may be problematic, it does not follow that there is anything wrong with someone possessing a given stereotypic trait (e.g., a gay man who likes fashion). Whether people fit or violate their group's stereotypes is immaterial to their value—we would hope that, rather than being judged or pressured based on the existence of a stereotype, people can be treated as individuals and judged on their own merit.

Conclusion

The gaydar myth is one example of Allport's observation that, "Groups that seem to *be* different will be thought (or made) to look different" (1954, p. 132, emphasis in original).

We found no support for the conclusion of prior work that the faces of gay men/lesbian women look different from the faces of straight people. Gay men and lesbian women's non-visible group status, however, is rendered purportedly visible by their cultural stereotypes. Relying on these stereotypes to infer orientation, however, is unlikely to yield accurate conclusions. Further, the folk concept of "gaydar" is a legitimizing myth that exacerbates this stereotyping process, covertly encouraging reliance on these stereotypes as categorization cues. Seeking out and emphasizing intergroup similarities rather than intergroup differences, combatting stereotypes, and dispelling legitimizing myths will help us progress toward the personal and societal goal of social equality.

Acknowledgments

We thank Lyn Abramson for comments on an earlier version of this article. We also thank our colleagues Markus Brauer, Judith Harackiewicz, Paula Niedenthal, Kristin Shutts, Steven Stroessner, and Carlie Allison, Sean Bennett, Emily Dix, Patrick Forscher, Luke Gray, Kristina Kellett, Greg Lyons, Chelsea Mitamura, Jessica Montag, Gary Oak, Jason Orne, Christopher Petsko, Magdalena Rychlowska, and Adrienne Wood for discussing the article and its ideas with us. Lastly, we thank Lauryn Besasie, Melanie Davies, Hannah Evans, Matt Huston, Nancy Lundin, Molly Maier, Kristen Merkitch, Amy Petermann, Alish Ridge-Montague, Peter Schaal, and Olivia Welk, as well as Tana Beck, Samantha Drake, Sonya Knudson, Renee Kramer, Shayla Krecklow, Kyle Kunz, Nancy Lundin, Megan Nelli, Katelyn Schuster, Shannon Smith, Ben Thomas, and Nicole Velez for their contributions to data collection, coding, and the preparation of this paper. The first and second authors were partially supported by NIH grants R01 GM111002 and R01 GM088477 during the collection of these data and the preparation of this paper.

References

Allport, GW. The nature of prejudice. Reading, MA: Addison-Wesley; 1954.

Ambady N, Hallahan M, Conner B. Accuracy of judgments of sexual orientation from thin slices of behavior. Journal of Personality and Social Psychology. 1999; 77:538. [PubMed: 10510507]

American Psychological Association. APA resolution on prejudice, stereotypes, and discrimination. 2006. Retrieved from http://www.apa.org/about/governance/council/policy/prejudice.pdf

Brebner JL, Martin D, Macrae CN. Dude looks like a lady: Exploring the malleability of person categorization. European Journal of Social Psychology. 2009; 39:1109–1119.

Bronski, M.; Pellegrini, A.; Amico, M. "You can tell just by looking" and 20 other myths about LGBT life and people. Boston: Beacon Press; 2013.

Cartei V, Reby D. Acting gay: Male actors shift the frequency components of their voices towards female values when playing homosexual characters. Journal of Nonverbal Behavior. 2012; 36:79–93.

Chapman L, Chapman J. Illusory correlations as an obstacle to the use of valid psychodiagnostic signs. Journal of Abnormal Psychology. 1969; 74:271–280. [PubMed: 4896551]

Cheryan S, Bodenhausen GV. When positive stereotypes threaten intellectual performance: The psychological hazards of 'model minority' status. Psychological Science. 2000; 11:399–402. [PubMed: 11228911]

Cox WTL, Abramson LY, Devine PG, Hollon SD. Stereotypes, prejudice, and depression: The integrated perspective. Perspectives on Psychological Science. 2012; 7:427–449. [PubMed: 26168502]

Cox WTL, Devine PG. Stereotyping to infer group membership creates plausible deniability for prejudice-based aggression. Psychological Science. 2014; 25:340–348. [PubMed: 24335602]

Czopp AM. When is a compliment not a compliment? Evaluating expressions of positive stereotypes. Journal of Experimental Social Psychology. 2008; 44:413–420.

DeBruine LM, Jones BC, Smith FG, Little AC. Are attractive men's faces masculine or feminine? The importance of controlling confounds in face stimuli. Journal of Experimental Psychology: Human Perception And Performance. 2010; 36:751–758. [PubMed: 20515201]

DeLamater JD, Hyde JS. Essentialism vs. social constructionism in the study of human sexuality. Journal of Sex Research. 1998; 35:10–18.

Dennis J. The boy who would be queen: Hints and closets on children's television. Journal of Homosexuality. 2009; 56:738–756. [PubMed: 19657933]

Devine PG. Stereotypes and prejudice: Their automatic and controlled components. Journal of Personality and Social Psychology. 1989; 56:5–18.

Devine PG, Forscher PS, Austin AJ, Cox WL. Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. Journal of Experimental Social Psychology. 2012; 48:1267–1278. [PubMed: 23524616]

Devine PG, Sherman S. Intuitive versus rational judgment and the role of stereotyping in the human condition: Kirk or Spock? Psychological Inquiry. 1992; 3:153–159.

Dotsch R, Wigboldus DHJ, van Knippenberg A. Biased allocation of faces to social categories. Journal of Personality and Social Psychology. 2011; 100:999–1014. [PubMed: 21443368]

Fiske, ST. Stereotyping, prejudice, and discrimination. In: Gilbert, DT.; Fiske, ST.; Lindzey, G., editors. The handbook of social psychology. 4th ed. Vol. 1 and 2. New York, NY: McGraw-Hill; 1998. p. 357-411.

Fiske ST. Venus and Mars or down to Earth: Stereotypes and realities of gender differences. Perspectives on Psychological Science. 2010; 5:688–692. [PubMed: 23678365]

Freeman JB, Johnson KL, Ambady N, Rule NO. Sexual orientation perception involves gendered facial cues. Personality and Social Psychology Bulletin. 2010; 36:1318–1331. [PubMed: 20682754]

Gaudio R. Sounding gay: Pitch properties in the speech of gay and straight men. American Speech. 1994; 69:30–57.

Gudelunas D. There's an App for that: The uses and gratifications of online social networks for gay men. Sexuality & Culture. 2012; 16:347–65.

Hamilton D, Gifford R. Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. Journal of Experimental Social Psychology. 1976; 12:392–407.

Hancock J, Toma C. Putting your best face forward: The accuracy of online dating photographs. Journal of Communication. 2009; 59:367–386.

Harry J. Sampling gay men. Journal of Sex Research. 1986; 22:21–34.

Haslam N, Rothschild L, Ernst D. Are essentialist beliefs associated with prejudice? British Journal of Social Psychology. 2002; 41:87–100. [PubMed: 11970776]

Hegarty P, Pratto F. The effects of social category norms and stereotypes on explanations for intergroup differences. Journal of Personality and Social Psychology. 2001; 80:723–735. [PubMed: 11374745]

Herek GM, Kimmel DC, Amaro H, Melton GB. Avoiding heterosexist bias in psychological research. American Psychologist. 1991; 44:957–963. [PubMed: 1958014]

Hyde JS. The gender similarities hypothesis. American Psychologist. 2005; 60:581–592. [PubMed: 16173891]

Inquisit 3.0.5.0 [Computer software]. Seattle, WA: Millisecond Software;

Jayaratne TE, Ybarra O, Sheldon JP, Brown TN, Feldbaum M, Pfeffer CA, Petty EM. White Americans' genetic lay theories of race differences and sexual orientation: Their relationship with prejudice toward blacks, and gay men and lesbians. Group Processes & Intergroup Relations. 2006; 9:77–94.

Johnson KL, Gill S, Reichman V, Tassinary LG. Swagger, sway, and sexuality: Judging sexual orientation from body motion and morphology. Journal of Personality and Social Psychology. 2007; 93:321–334. [PubMed: 17723051]

Judd CM, Park B. Definition and assessment of accuracy in social stereotypes. Psychological Review. 1993; 100:109–128. [PubMed: 8426877]

Kahneman D, Tversky A. On the psychology of prediction. Psychological Review. 1973; 80:237–251.

Kay AC, Day MV, Zanna MP, Nussbaum A. The insidious (and ironic) effects of positive stereotypes. Journal Of Experimental Social Psychology. 2013; 49:287–291.

Kite ME, Deaux K. Gender belief systems: Homosexuality and the implicit inversion theory. Psychology of Women Quarterly. 1987; 11:83–96.

Lyons M, Lynch A, Brewer G, Bruno D. Detection of sexual orientation ("gaydar") by homosexual and heterosexual women. Archives of Sexual Behavior. 2014; 43:345–352. [PubMed: 23813041]

Matthews C, Hill C. Gay until proven straight: Perceptions of male interior designers from male practitioner and student perspectives. Journal of Interior Design. 2011; 36:15–34.

Mitroff SR, Biggs AT. The ultra-rare-item effect: Visual search for exceedingly rare items is highly susceptible to error. Psychological Science. 2014; 25:284–289. [PubMed: 24270463]

Nadal, KL. That's so gay! Microaggressions and the lesbian, gay, bisexual, and transgender community. Washington, DC US: American Psychological Association; 2013.

Nicholas CL. Gaydar: Eye-gaze as identity recognition among gay men and lesbians. Sexuality and Culture. 2004; 8:60–86.

Nisbett RE, Zukier H, Lemley RE. The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. Cognitive Psychology. 1981; 13:248–277.

Olivola CY, Funk F, Todorov A. Social attributions from faces bias human choices. Trends in Cognitive Sciences. 2014; 18:566–570. [PubMed: 25344029]

Olivola CY, Todorov A. Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. Journal of Experimental Social Psychology. 2010; 46:315–324.

Pettigrew TF. The ultimate attribution error: Extending Allport's cognitive analysis of prejudice. Personality and Social Psychology Bulletin. 1979; 5:461–476.

Plöderl M. Bayesian advice for gaydar-based picking up: Commentary on Lyons, Lynch, Brewer, and Bruno (2013). Archives of Sexual Behavior. 2014; 43:7–9. [PubMed: 23942737]

Pratto F, Sidanius J, Stallworth LM, Malle BF. Social dominance orientation: A personality variable predicting social and political attitudes. Journal of Personality and Social Psychology. 1994; 67:741–763.

Prentice DA, Miller DT. Psychological essentialism of social categories. Current Directions in Psychological Science. 2007; 16:202–206.

Reuter, DF. Gaydar: The ultimate insider guide to the gay sixth sense. NY: Crown; 2002.

Rieger G, Linsenmeier JA, Gygax L, Garcia S, Bailey JM. Dissecting "gaydar": Accuracy and the role of masculinity-femininity. Archives of Sexual Behavior. 2010; 39:124–140. [PubMed: 18810629]

Riggs, M. Ethnic Notions [Motion picture]. United States: California Newsreel; 1987. (Producer, Director).

Rothbart, M.; Taylor, M. Category labels and social reality: Do we view social categories as natural kinds?. In: Semin, GR.; Fiedler, K., editors. Language, interaction and social cognition. Thousand Oaks, CA US: Sage Publications, Inc.; 1992. p. 11-36.

Rule NO, Ambady N. Brief exposures: Male sexual orientation is accurately perceived at 50 ms. Journal of Experimental Social Psychology. 2008; 44:1100–1105.

Rule NO, Ambady N, Adams RB Jr, Macrae CN. Accuracy and awareness in the perception and categorization of male sexual orientation. Journal of Personality and Social Psychology. 2008; 95:1019–1028. [PubMed: 18954191]

Rule NO, Ambady N, Hallett KC. Female sexual orientation is perceived accurately, rapidly, and automatically from the face and its features. Journal of Experimental Social Psychology. 2009; 45:1245–1251.

Rule NO, Ishii K, Ambady N, Rosen KS, Hallett KC. Found in translation: Cross-cultural consensus in the accurate categorization of male sexual orientation. Personality and Social Psychology Bulletin. 2011; 37:1449–1507. [PubMed: 21653580]

Rule NO, Macrae CN, Ambady N. Ambiguous group membership is extracted automatically from faces. Psychological Science. 2009; 20:441–443. [PubMed: 19399971]

Russo, V. The celluloid closet: Homosexuality in the movies. New York, NY: Arbor House Publishing; 1987.

Savin-Williams RC. Who's gay? Does it matter? Current Directions in Psychological Science. 2006; 15:40–44.

Seinfeld, JA. David, L. quote from: Charles, L. (Writer) & Cherones, T. (Director) (1993). The outing [Television series episode]. In David, L. (Executive Producer), Seinfeld. New York, NY: NBC Corporation; 1993.

Shelp SG. Gaydar: Visual detection of sexual orientation among gay and straight men. Journal of Homosexuality. 2002; 44:1–14. [PubMed: 12856753]

Smyth R, Jacobs G, Rogers H. Male voices and perceived sexual orientation: An experimental and theoretical approach. Language in Society. 2003; 32:329–350.

Steele CM. A threat in the air: How stereotypes shape intellectual identity and performance. American Psychologist. 1997; 52:613–629. [PubMed: 9174398]

Substance Abuse and Health Services. Behavioral Health, United States, 2012. Substance Abuse and Mental Health Services Administration, (13-4797). 2013. Retrieved from: http://www.samhsa.gov/data/2012BehavioralHealthUS/2012-BHUS.pdf

Tabak JA, Zayas V. The roles of featural and configural face processing in snap judgments of sexual orientation. PLoS ONE. 2012; 7:e36671. [PubMed: 22629321]

Todorov A, Porter JM. Misleading first impressions different for different facial images of the same person. Psychological Science. 2014; 25:1404–1417. [PubMed: 24866921]

Toma C, Hancock J, Ellison N. Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles. Personality and Social Psychology Bulletin. 2008; 34:1023–1036. [PubMed: 18593866]

Uhlmann E, Cohen GL. 'I think it, therefore it's true': Effects of self-perceived objectivity on hiring discrimination. Organizational Behavior And Human Decision Processes. 2007; 104:207–223.

Van Borsel J, Van de Putte A. Lisping and male homosexuality. Archives of Sexual Behavior. 2014; 43:1159–1163. [PubMed: 24578106]

Appendix - Signal Detection Statistics

	Gay-Stereotypic Trials				**Stereotype-Neutral Trials**				**Straight-Stereotypic Trials**			
	d-prime		**c**		**d-prime**		**c**		**d-prime**		**c**	
Study 1A	−0.02	(0.5)	−0.05	(0.5)	0.03	(0.5)	0.39*	(0.4)	0.06	(0.5)	0.45*	(0.4)
Study 1B	0.12	(0.4)	−0.06	(0.5)	0.09	(0.5)	0.33*	(0.4)	−0.07	(0.4)	0.37*	(0.4)
Study 2A	0.14*	(0.5)	−0.20*	(0.4)	0.19*	(0.4)	0.30*	(0.4)	0.21*	(0.4)	0.30*	(0.5)
Study 2B	0.22*	(0.5)	−0.20*	(0.4)	0.28*	(0.4)	0.28*	(0.4)	0.23*	(0.5)	0.21*	(0.5)
	Quality-Unmatched				Quality-Matched							
	d-prime		c		d-prime		c					
Study 3	0.12*	(0.3)	0.45*	(0.4)	−0.00	(0.3)	0.32*	(0.4)				
Study 4	0.21*	(0.2)	0.43*	(0.5)	−0.07	(0.2)	0.66*	(0.6)				

Note. Means and standard deviations of signal detection statistics for Studies 1–4. For each study, we coded the self-identified gay men's and lesbian women's pictures as "signal present" and the self-identified straight men's and straight women's pictures as "signal absent."

All means that are significantly different from 0 via one-sample t-test ($p < 0.05$) are marked with an asterisk. Positive, significant d-prime statistics indicate that participants rated the gay men's/lesbians' pictures as gay/lesbian more often than the straight men's/women's pictures. Positive, significant c statistics indicate that, relative to chance, participants were overall biased towards guessing "straight," whereas negative c statistics indicate that they were biased in favor of guessing "gay" or "lesbian." Participants could only distinguish orientation from the pictures when they were not matched on quality, as in Study 2 and the Quality-Unmatched conditions of Studies 3 and 4. When the stimuli were matched for quality, no effect of picture arose.

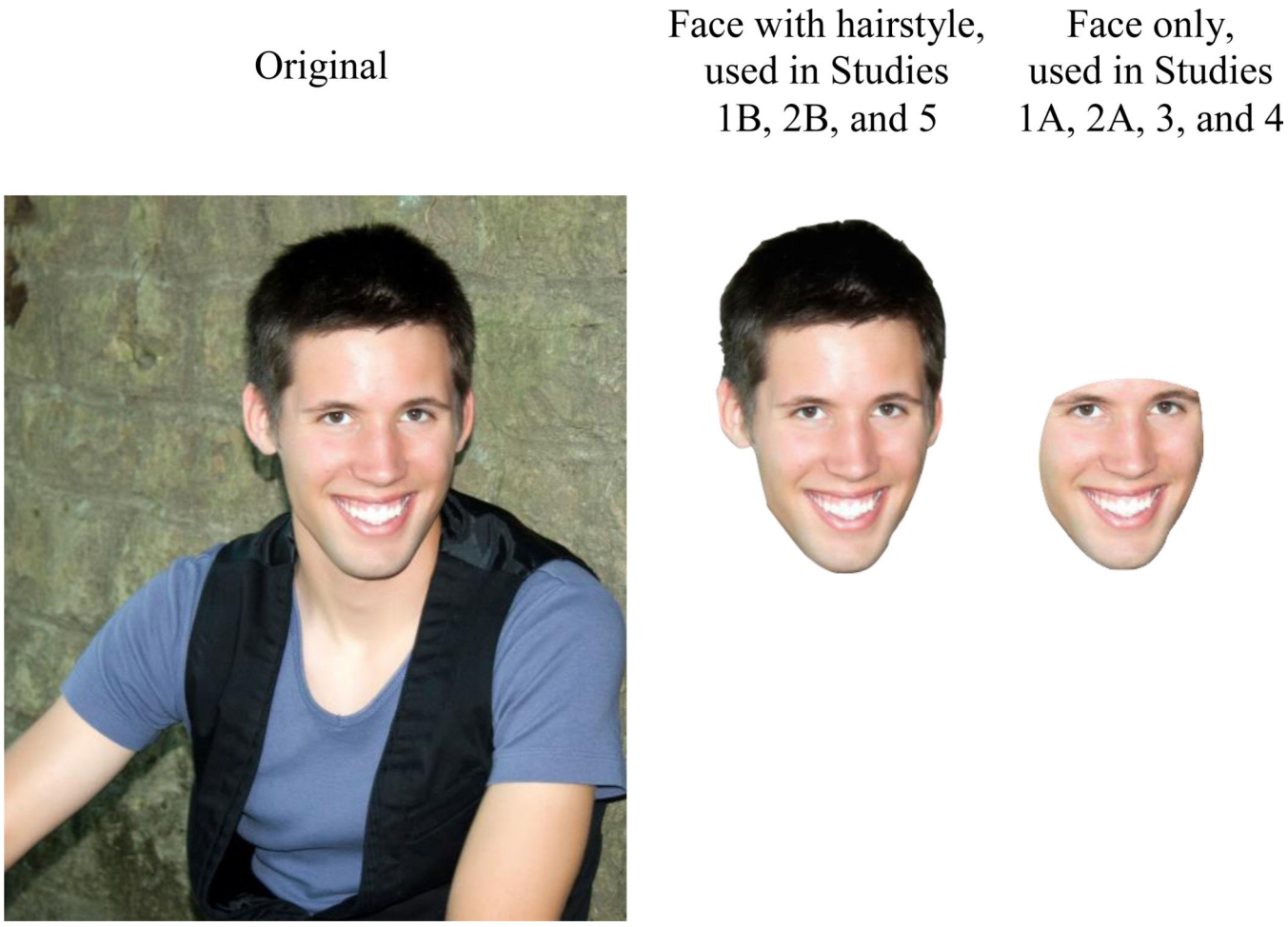

Figure 1.
Cropping example. This is a picture of a lab volunteer, cropped like the stimuli used in the present studies. The cropped pictures were placed on a white background larger than is shown in this figure.

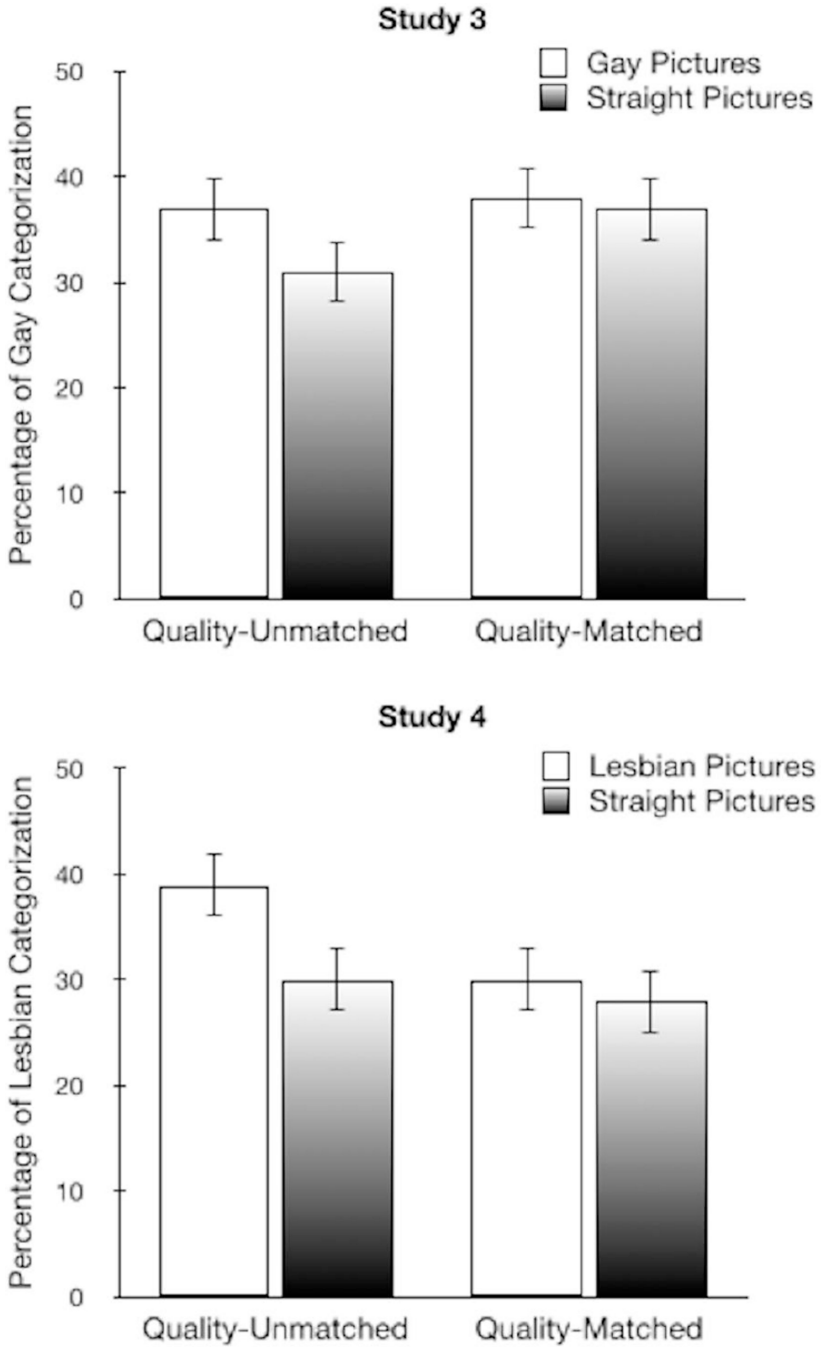

Figure 2.

Study 3 (Top) and Study 4 (Bottom) Results. Error bars are the standard error of the means. For both male (Study 3) and female (Study 4) targets, perceivers could not distinguish orientation from the pictures when the stimuli have been matched for quality.

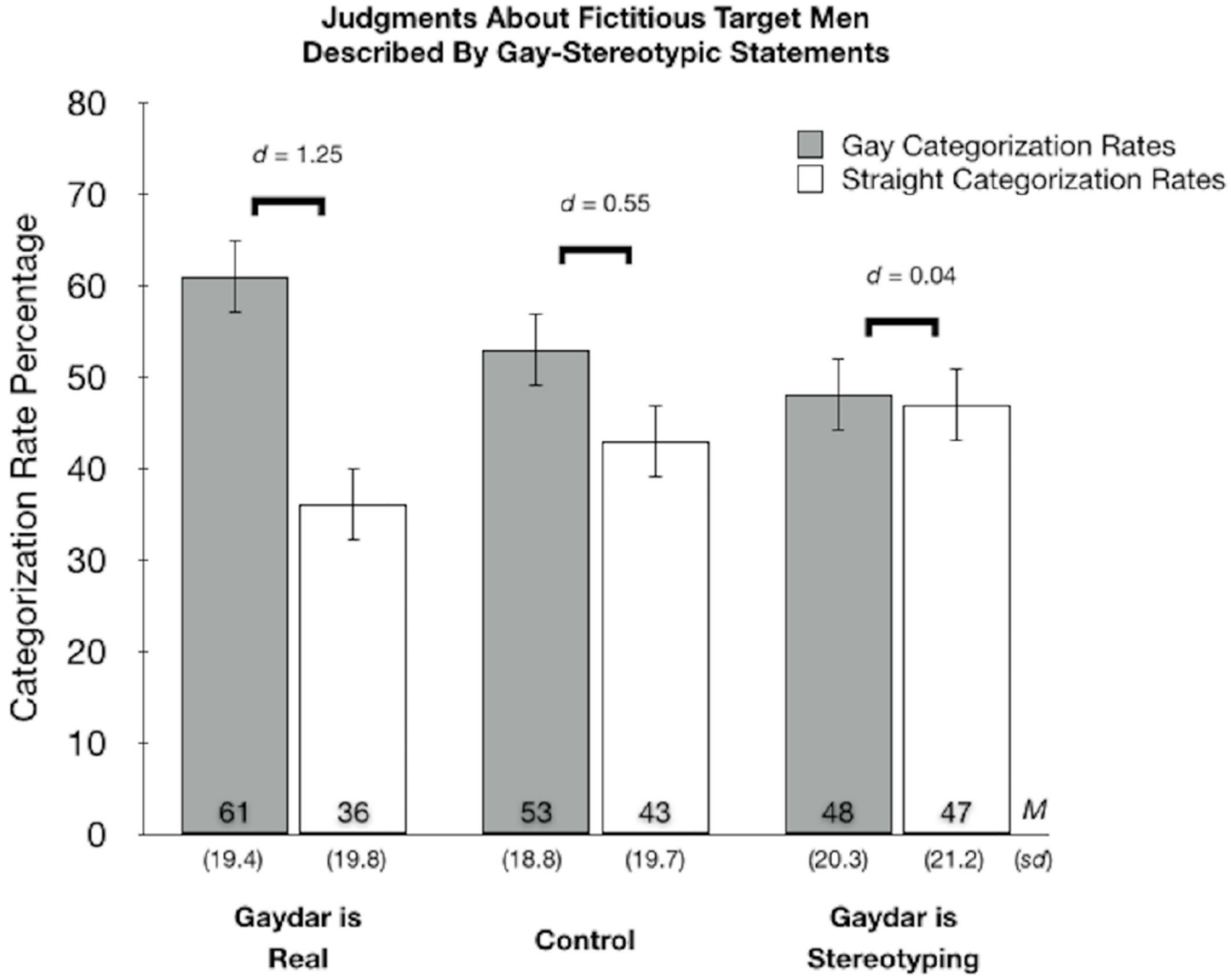

Figure 3.

Study 5 Results. Error bars are the standard error of the means. A 2 (Categorization Rate Type: Straight vs. Gay) X 3 (Gaydar Belief condition: Gaydar is Real vs. Control vs. Gaydar is Stereotyping) mixed ANOVA revealed that the main effect of categorization rate type, F (1, 230) = 22.694, $p < 0.001$, $\eta^2 = 0.085$, was qualified by an interaction with Gaydar Belief condition, F (2, 230) = 7.508, $p = 0.001$, $\eta^2 = 0.056$. After being told that gaydar is real, participants guessed that gay-stereotypic men were gay much more often than straight, and this difference was much larger than for participants in the control group. When participants were told that gaydar is merely another label for stereotyping, they were equally likely to guess that a gay-stereotypic man was gay or straight.

Figure 4.

Fallacies of Accurate Stereotype-Based Gaydar. Mistaken conclusions about the accuracy of stereotype-based gaydar are facilitated by the *base rate fallacy* and the *fallacy of the converse*. Evidence of an accurate Group → Attribute stereotype (e.g., gay men possess attribute X), as in the top panel, must be adjusted for the base rate of gay men in the population (middle panel), before we can estimate the accuracy of the converse Attribute → Group stereotype (e.g., men who possess attribute X are gay), as in the bottom panel. Even implausibly high levels of Group → Attribute accuracy, as pictured, are highly unlikely to yield better-than-chance Attribute → Group accuracy. Mathematically, accurate stereotype-based gaydar seems highly infeasible.

Table 1

Gay-Stereotypic, Stereotype Neutral, and Straight-Stereotypic Statements

Statement	*M*	*SD*	Statement	*M*	*SD*	Statement	*M*	*SD*
Gay-Stereotypic	**3.09**	**(0.387)**	**Stereotype-Neutral**	**4.00**	**(0.055)**	**Straight-Stereotypic**	**5.15**	**(0.240)**
He is a hair dresser	2.39	(0.766)	He likes to travel	4.00	(0.338)	He plays football	5.61	(0.934)
He is an interior designer	2.53	(0.774)	He likes children	4.00	(0.487)	He hunts	5.58	(0.937)
He is a Cher fan	2.61	(0.838)	He is outgoing	3.97	(0.654)	He watches Sports Center	5.31	(0.786)
He is fashionable	2.92	(0.732)	He likes spaghetti	3.97	(0.167)	He drives a pick-up truck	5.31	(0.889)
He likes going to the mall	2.93	(0.997)	He likes animals	3.97	(0.506)	He watches NASCAR	5.31	(1.009)
He likes musical theater	3.03	(0.654)	He has two siblings	3.97	(0.167)	He does not care about fashion	5.11	(0.747)
He likes shopping	3.06	(0.791)	He likes listening to music	3.97	(0.291)	He is a firefighter	5.06	(0.924)
He has a stylish home	3.08	(0.841)	He has a good sense of humor	4.03	(0.291)	He is a rapper	5.06	(0.984)
He does not like sports	3.31	(0.668)	He has an accent	4.03	(0.506)	He is a police officer	4.97	(0.845)
He is a nurse	3.39	(0.599)	He likes to drink alcohol	4.06	(0.232)	He likes cars	4.97	(0.845)
He spends his time working for equal rights	3.44	(0.652)	He likes watching movies	3.94	(0.333)	He rides a motorcycle	4.97	(0.774)
He likes to be well dressed	3.5	(0.609)	He likes to read	3.94	(0.232)	He is good at sports	4.97	(1.000)
He likes dancing	3.53	(0.654)	He likes to go jogging	3.94	(0.410)	He likes to play basketball for fun	4.94	(0.954)
He likes to be well groomed	3.58	(0.604)	He is active	4.14	(0.424)	He is a Republican	4.86	(1.089)
$-1.16 \leq skews \leq 0.13$			$-4.05 \leq skews \leq 4.05$			$-0.37 \leq skews \leq 0.68$		
$-1.37 \leq kurtoses \leq 0.56$			$-1.85 \leq kurtoses \leq 15.26$			$-1.37 \leq kurtoses \leq 0.56$		
$t\,(13) = -8.751$, $p < 0.001$			$t\,(13) = -0.337$, $p = 0.741$			$t\,(13) = 17.854$, $p < 0.001$		

Note. Descriptive statistics from the pretesting of the selected stimulus statements. Ratings were on a scale from 1 (*very gay*) to 7 (*very straight*). To verify our selections, the one-sample t-tests compared the set of statements to the scale midpoint, 4 (*neither gay nor straight*).

J Sex Res. Author manuscript; available in PMC 2016 February 01.

Table 2

Means and standard deviations for Studies 1-5

	Study 1A						Study 1B					
Statements	**Gay Pics**		**Straight Pics**		**All Pics**		**Gay Pics**		**Straight Pics**		**All Pics**	
Gay-Stereotypic	57%	(28.6)	56%	(24.8)	57%	(23.6)	60%	(26.7)	56%	(26.5)	58%	(25.4)
Stereotype-Neutral	27%	(21.3)	27%	(21.3)	27%	(18.2)	33%	(22.5)	32%	(21.1)	32%	(17.5)
Straight-Stereotypic	21%	(20.9)	20%	(23.1)	20%	(19.7)	27%	(21.5)	26%	(22.6)	27%	(19.1)
All Statements	35%	(15.0)	34%	(14.7)	35%	(13.3)	40%	(14.6)	38%	(15.6)	39%	(13.5)
	Study 2A						Study 2B					
Statements	Gay Pics		Straight Pics		All Pics		Gay Pics		Straight Pics		All Pics	
Gay-Stereotypic	65%	(24.6)	57%	(24.8)	61%	(21.6)	66%	(22.4)	55%	(23.1)	61%	(19.4)
Stereotype-Neutral	42%	(22.7)	31%	(19.8)	36%	(17.3)	42%	(23.9)	29%	(18.8)	35%	(17.6)
Straight-Stereotypic	36%	(25.7)	25%	(22.1)	30%	(21.4)	42%	(24.8)	28%	(24.8)	35%	(21.4)
All Statements	48%	(14.3)	37%	(12.5)	42%	(11.7)	50%	(13.5)	37%	(12.6)	44%	(10.8)
	Study 3				Study 4							
Condition	Gay Pics		Straight Pics		Lesbian Pics		Straight Pics					
Quality-Matched	38%	(18.3)	37%	(16.9)	30%	(17.6)	28%	(18.1)				
Quality-Unmatched	37%	(17.7)	31%	(14.6)	39%	(17.2)	30%	(18.6)				
	Study 5											
Conditions:	**Gaydar is Real**				**Control**				**Gaydar is Stereotyping**			
Statements	Gay Judgments		Straight Judgments		Gay Judgments		Straight Judgments		Gay Judgments		Straight Judgments	
Gay-Stereotypic	61%	(19.4)	36%	(19.8)	53%	(18.8)	43%	(19.7)	48%	(20.3)	47%	(21.2)
Stereotype-Neutral	35%	(14.5)	57%	(18.0)	36%	(17.0)	58%	(17.3)	33%	(15.2)	60%	(18.6)
Straight-Stereotypic	29%	(20.1)	66%	(20.8)	35%	(22.6)	62%	(22.3)	33%	(19.2)	63%	(19.4)
All Statements	42%	(10.3)	53%	(12.0)	41%	(12.2)	54%	(11.5)	38%	(10.4)	57%	(11.5)

Note. For Studies 1–4, percentages of gay categorization. For Study 5, percentages of gay categorization and straight categorization.