

## Straight Until Proven Gay: A Systematic Bias Toward Straight Categorizations in Sexual Orientation Judgments

David J. Lick  
New York University

Kerri L. Johnson  
University of California, Los Angeles

Perceivers achieve above chance accuracy judging others' sexual orientations, but they also exhibit a notable response bias by categorizing most targets as straight rather than gay. Although a *straight categorization bias* is evident in many published reports, it has never been the focus of systematic inquiry. The current studies therefore document this bias and test the mechanisms that produce it. Studies 1–3 revealed the straight categorization bias cannot be explained entirely by perceivers' attempts to match categorizations to the number of gay targets in a stimulus set. Although perceivers were somewhat sensitive to base rate information, their tendency to categorize targets as straight persisted when they believed each target had a 50% chance of being gay (Study 1), received explicit information about the base rate of gay targets in a stimulus set (Study 2), and encountered stimulus sets with varying base rates of gay targets (Study 3). The remaining studies tested an alternate mechanism for the bias based upon perceivers' use of gender heuristics when judging sexual orientation. Specifically, Study 4 revealed the range of gendered cues compelling gay judgments is smaller than the range of gendered cues compelling straight judgments despite participants' acknowledgment of equal base rates for gay and straight targets. Study 5 highlighted perceptual experience as a cause of this imbalance: Exposing perceivers to hyper-gendered faces (e.g., masculine men) expanded the range of gendered cues compelling gay categorizations. Study 6 linked this observation to our initial studies by demonstrating that visual exposure to hyper-gendered faces reduced the magnitude of the straight categorization bias. Collectively, these studies provide systematic evidence of a response bias in sexual orientation categorization and offer new insights into the mechanisms that produce it.

*Keywords:* social perception, sexual orientation perception, gaydar, response bias

People possess a remarkable capacity to judge others according to their social category memberships, exhibiting near-perfect accuracy when categorizing identities such as race or sex (Bodenhause & Peery, 2009; Johnston, Miles, & Macrae, 2010). Such high levels of accuracy are unsurprising, given that cues to race and sex are unambiguously etched in the face and body (Johnson & Tassinari, 2005; Lick, Johnson, & Gill, 2013; Martin & Macrae, 2007). More intriguing is the fact that perceivers also achieve above-chance accuracy when categorizing more visually ambigu-

ous social identities, including religious identification, political affiliation, and sexual orientation (Tskhay & Rule, 2013). Of these more ambiguous identities, sexual orientation dominate the research literature, with twice the number of studies investigating sexual orientation compared to all other ambiguous social category dimensions combined (Tskhay & Rule, 2013). This work has consistently revealed that perceivers quickly and accurately categorize others' sexual orientations based upon minimal exposure, leading to the conclusion that perceivers possess a highly sensitive "gaydar" (Tabak & Zayas, 2012).

Despite these claims, a recent meta-analysis revealed the average effect size for accuracy in sexual orientation judgments is quite modest ( $M_{effect\ size} = 0.29$ ; Tskhay & Rule, 2013). This is not to say that accuracy is unimportant, but rather that perceivers' ability to categorize sexual orientation is hardly error-free. As such, scientific and popular writing focused solely on accuracy risks overlooking important aspects of sexual orientation categorization. Considering the means by which observers achieve accurate sexual orientation judgments further underscores this point: Whereas observers tend to correctly categorize the sexual orientation of straight men and women, their categorizations of gay men and lesbians are decidedly more error-prone (Freeman, Johnson, Ambady, & Rule, 2010; Johnson, Gill, Reichman, & Tassinari, 2007;

David J. Lick, Department of Psychology, New York University; Kerri L. Johnson, Departments of Psychology and Communication Studies, University of California, Los Angeles.

This research was supported by a National Science Foundation Graduate Research Fellowship and Eugene V. Cota-Robles Fellowship (David J. Lick). We thank Jon Freeman for stimuli, participants of the 2013 Duck Conference on Social Cognition for thoughtful feedback on a subset of these studies, and members of the Social Communication Lab for their assistance with data collection and comments on prior versions of this article.

Correspondence concerning this article should be addressed to David J. Lick, Department of Psychology, New York University, New York, NY 10003. E-mail: dl136@nyu.edu

Linville, 1998). In fact, a cursory review of existing literature reveals that perceivers categorize the majority of targets they encounter as straight rather than gay. The near-exclusive focus on accuracy in prior research has masked this *straight categorization bias*, providing an incomplete picture of sexual orientation categorization. The current studies systematically investigate the straight categorization bias to provide a more comprehensive analysis of the outcomes resulting from this consequential form of social judgment.

### Sexual Orientation Categorization

Social categorization—the process of discerning others' social identities—is a foundational aspect of human psychology that carries serious interpersonal consequences (Allport, 1954). For example, categorizing others as members of stigmatized groups can elicit stereotypes and compel prejudice (Bodenhausen & Macrae, 1998; Devine, 1989; Wilder, 1986). These are matters of serious concern for sexual minority individuals. Indeed, whereas hate crimes targeting many stigmatized groups have declined in recent years, hate crimes targeting sexual minorities have increased (Federal Bureau of Investigation, 2011). In fact, gay men and lesbians experience some of the highest rates of prejudice known today (Katz-Wise & Hyde, 2012), which have been linked to health deficits ranging from psychopathology to cardiovascular disease, cancer, and suicide (Hatzenbuehler, 2009; Lick, Durso, & Johnson, 2013; Meyer, 2003).

The consequences described above likely contribute to burgeoning scientific interest in sexual orientation categorization. Research on this topic has accelerated over the past decade, with results converging upon several conclusions. First, sexual orientation categorizations are surprisingly accurate. Even when based upon minimal information, perceivers accurately judge men's and women's sexual orientations at above-chance levels (Ambady, Hallahan, & Conner, 1999; Johnson et al., 2007; Lick et al., 2013; Linville, 1998; Munson & Babel, 2007; Rendall, Vasey, & McKenzie, 2008; Rule & Ambady, 2008; Rule, Ambady, Adams, & Macrae, 2008; Rule, Ambady, & Hallett, 2009; Stern, West, Jost, & Rule, 2013). This pattern is observed for judgments of isolated facial cues (Rule et al., 2008), vocal samples (Linville, 1998; Rendall et al., 2008), and body features (Ambady et al., 1999; Johnson et al., 2007; Lick et al., 2013), giving rise to modest but reliable accuracy in sexual orientation judgments overall (Rieger, Linsenmeier, Gygax, Garcia, & Bailey, 2010; Tskhay & Rule, 2013).

Second, perceivers endorse gender inversion stereotypes related to sexual minorities, and they use these stereotypes as heuristics for inferring unknown others' sexual orientations (Kite & Deaux, 1987). Specifically, perceivers tend to judge normatively gendered targets (masculine men, feminine women) as straight but more androgynous targets (feminine men, masculine women) as gay. Use of this gender inversion heuristic for categorizing sexual orientation is robust, persisting across judgments drawn from facial features (Freeman et al., 2010), body cues (Johnson et al., 2007), and vocal characteristics (Gaudio, 1994).

Third, the gendered cues perceivers use to judge sexual orientation guide broader social evaluations. For instance, androgyny tends to be evaluated negatively in general (Johnson & Tassinari, 2007) and among sexual minorities in particular (Lick & Johnson,

2014a).<sup>1</sup> Moreover, androgyny leads observers to believe targets are flaunting their sexuality, in turn prompting prejudice (Lick, Johnson, & Gill, 2014). Finally, the androgynous cues that compel sexual minority categorizations require cognitive effort to process, and such disfluency arouses negative evaluations (Lick & Johnson, 2013). Thus, gay men and lesbians experience prejudice in part because of the very cues perceivers use to determine their sexual orientation.

### Decomposing the Outcomes of Sexual Orientation Categorization

Although researchers have begun probing the perceptual mechanisms and evaluative consequences of sexual orientation categorization, the vast majority of work in this area has focused on categorization accuracy. Given the field's primary focus on accuracy, it is surprising to note that many published reports present a simplified version of the findings. In particular, scientists and laypeople alike tend to combine gay and straight category judgments when estimating the accuracy of sexual orientation categorizations. This tendency is problematic because it obscures the decision strategies underlying perceivers' judgments. Consider, for example, the recent estimate that sexual orientation categorizations are 64% accurate overall (Tskhay & Rule, 2013). When described in aggregate, it is impossible to discern how observers achieved such accuracy. The statistic could reflect comparable success when judging gay and straight targets (e.g., correctly categorizing 32/50 gay targets and 32/50 straight targets), but it could also reflect an imbalance (e.g., correctly categorizing 47/50 straight targets and 17/50 gay targets). Knowing whether judgments are made in a balanced or biased manner would provide a more complete understanding of sexual orientation categorization.

Signal detection theory provides an ideal means of quantifying the outcomes of sexual orientation judgments (Stanislaw & Todorov, 1999). Within this framework, categorizing targets as gay or straight can result in one of four outcomes: gay targets are correctly judged to be gay (*hit*) or incorrectly judged to be straight (*miss*); straight targets are correctly judged to be straight (*correct rejection*) or incorrectly judged to be gay (*false alarm*). The relative proportions of these four outcomes can be used to estimate two parameters of theoretical interest—sensitivity and bias. Sensitivity indicates the extent to which perceivers can discern the category in question (here, the category gay).<sup>2</sup> Bias indicates the extent to which perceivers favor one category over another in their decision-making. If the ratio of correct rejections to false alarms is

<sup>1</sup> Prior literature in this area has referred to the relatively feminine features common among gay men and the relatively masculine features common among lesbian women as *gender-atypical*. As the editor pointed out, however, the word "atypical" could capture deviations from the norm in either direction (e.g., very masculine men or very feminine men). To avoid confusion, we instead use the term androgynous to describe masculine men and feminine women.

<sup>2</sup> In this example, we arbitrarily defined gay targets as signal and straight targets as noise. Doing so is theoretically justified, insofar as perceivers often consider heterosexuality to be normative and homosexuality to be a departure from that norm (Bem, 1993). Moreover, this approach is common to almost all published studies of sexual orientation categorization. Still, it is worth noting that this convention is arbitrary; we could have defined straight targets as signal, which would reverse the numeric pattern of results but maintain a similar interpretation.

less than the ratio of hits to misses, perceivers exhibit a biased tendency to categorize targets as gay rather than straight. If the ratio of correct rejections to false alarms is greater than the ratio of hits to misses, perceivers exhibit a biased tendency to categorize targets as straight rather than gay. Although these strategies can achieve equivalent levels of sensitivity, they do so by different means. The former prioritizes categorizing gay targets as gay whereas the latter prioritizes categorizing straight targets as straight.

Because of the level of detail it provides, signal detection has become the gold standard in studies of social categorization. Nevertheless, a majority of published research has focused exclusively on the sensitivity of sexual orientation judgments, with little consideration of response bias. Such a narrow focus risks providing an incomplete or misleading view of sexual orientation categorization. Here, we begin exploring whether perceivers use a balanced or biased decision strategy when judging unknown others' sexual orientations.

### An Underappreciated Bias in Sexual Orientation Judgments

Although generally not scrutinized, several studies have reported signal detection parameters relevant to our question. The available data are highly consistent, indicating a strong bias toward straight categorizations in thin slice judgments of sexual orientation. For example, one study found that although perceivers were sensitive to sexual orientation cues in dynamic body motions, they also displayed a notable response bias toward straight categorizations (Johnson et al., 2007, Study 3). Later studies revealed similar biases in sexual orientation judgments drawn from facial features (Freeman et al., 2010). This straight categorization bias is extensive, occurring for judgments of faces belonging to women (Rule et al., 2008), men (Rule, Ishii, Ambady, Rosen, & Hallett, 2011), racial minorities (Johnson & Ghavami, 2011), and cultural outgroups (Rule et al., 2011). A similar bias emerges for judgments that rely on cues in modalities other than vision (e.g., voice; Linville, 1998). Thus, there appears to be a pronounced bias favoring straight judgments in published research on sexual orientation categorization.

Although this bias is consistent across studies, it has received almost no empirical scrutiny. Instead, response bias has historically been calculated only in the service of quantifying sensitivity, with bias parameters relegated to footnotes and supplements if mentioned at all. Still, we contend that a thorough investigation of response bias can offer new insights into the decision-making processes underlying sexual orientation categorization. The current studies offer the first systematic investigation of response bias in sexual orientation categorization, clarifying whether and why perceivers categorize most targets they encounter as straight rather than gay. Studies 1–3 use various paradigms to document the straight categorization bias and test one possible explanation for it—namely, that perceivers tend to categorize unknown others as straight because of the low base rate of sexual minorities in the population. Studies 4–6 test another explanation based on the gender inversion heuristic many perceivers use to categorize sexual orientation. By documenting both the presence and causes of the straight categorization bias, our work seeks to provide new information

about the judgment strategies people use when confronted with the task of discerning perceptually ambiguous identities on the basis of minimal information.

### Study 1

The straight categorization bias could arise through several distinct mechanisms. Studies 1–3 tested whether it reflects perceivers' appreciation of the low population base rate of sexual minorities. Perceivers often attempt to probability match by categorizing a class of stimuli proportional to the frequency with which it exists in the population (Erev & Barron, 2005; Estes, 1976). This tendency is relevant here because lesbians and gay men make up about 3.5% of the U.S. population (Gates, 2011), whereas most studies of sexual orientation categorization have included equal numbers of gay and straight targets without informing participants. The straight categorization bias might therefore emerge because perceivers consult population base rates when the proportion of gay to straight targets in a stimulus set is unknown.

Although feasible, there is reason to believe that base rate matching provides an incomplete explanation for the straight categorization bias. In particular, prior research has shown that observers tend to underutilize base rate information relative to individuating information when making probabilistic judgments (Bar-Hillel, 1980; Tversky & Kahneman, 1981). This tendency even applies to social categorizations for which base rates are made obvious. For example, Olivola and Todorov (2010, Study 1) found that participants underutilized explicit base rate information when categorizing Americans' political party affiliations, relying more heavily on individuating information supplied by facial cues. While these findings suggest that base rate matching is unlikely to provide a full explanation for the straight categorization bias, strong conclusions await a direct test of the hypothesis in the domain of sexual orientation categorization.

One method of testing whether base rate matching accounts for the biased tendency to categorize unknown others as straight would be to examine response outcomes when decisions are statistically independent of one another. If the straight categorization bias is driven entirely by perceivers' attempts to match their categorizations to the true base rate of gay targets, there should be no bias when the exact probability of a target being gay is specified as 50% on a trial-by-trial basis. Study 1 explored this possibility.

### Method

**Participants.** Sixty Internet users completed the study ( $M_{age} = 41.68$ ; 62% women; 72% White; 90% straight). In this and all subsequent studies, excluding participants who identified as lesbian, gay, or bisexual did not alter the direction or significance of any results, so we retained their responses. Although we had no reason to believe response bias would differ across male and female participants, we also examined moderating effects of perceiver sex in all studies. None of these effects reached statistical significance, so we do not discuss them further.

**Stimuli.** Stimuli were a subset of images used in prior studies of sexual orientation categorization (Freeman et al., 2010). The selected images were 40 grayscale facial photographs of real people who varied by sex and sexual orientation (10 gay men, 10 straight men, 10 lesbian women, and 10 straight women). All

photographs were standardized to  $200 \times 288$  pixels and depicted White individuals devoid of facial hair, jewelry, and visible tattoos.

**Procedure.** We recruited Mechanical Turk users for a study about their impressions of other people, with no mention of base rates or sexual orientation. After providing consent, participants were redirected to the survey hosting website Qualtrics, where they were instructed to categorize strangers' sexual orientations on the basis of facial photographs. Critically, these categorizations were embedded within a game called "Two Doors." On each trial, participants were presented images of two doors—one labeled Door A and the other labeled Door B. They were informed that one door concealed a picture of a straight person and the other door concealed a picture of a gay person, with the location of the gay/straight faces randomized across trials. On each trial, participants chose a door and guessed the sexual orientation of the face it revealed. In this way, each trial was statistically independent and had a 50% chance of revealing a gay face.

Programming constraints precluded a fully randomized presentation order. To overcome this limitation, we created a fixed random order for the faces using a random number generator (forward condition), and then we reversed it (backward condition). Participants were randomly assigned to one of these two presentation orders, viewing the 40 faces in the specified order regardless of which door they selected on each trial.

Participants were given unlimited time to categorize each face and they received no feedback about the accuracy of their judgments. After completing all 40 trials, participants responded to a manipulation check testing whether they understood the statistical constraints of the design: "Based on the information provided at the beginning of the study, what percentage of the people you saw in this study were actually gay?" Finally, participants reported demographic information before being debriefed.

## Results and Discussion

If base rate matching fully explains the straight categorization bias observed in prior work, then participants should show no bias when they believed each trial had a 50% chance of revealing a gay face. We conducted signal detection analyses to estimate the magnitude of response bias within each condition. Specifically, we coded correct gay categorizations as hits and correct straight categorizations as correct rejections, computing response bias ( $c$ ) using standard algorithms (Stanislaw & Todorov, 1999). In this coding scheme,  $c$  values of 0 indicate no bias, positive  $c$  values indicate a conservative bias (i.e., tendency to apply straight category labels), and negative  $c$  values indicate a liberal bias (i.e., tendency to apply gay category labels).

**Manipulation check.** We first analyzed responses to the manipulation check to ensure that participants understood the method. On average, participants reported that 46.80% of the faces they saw were gay ( $SD = 19.10$ ). A one-sample  $t$  test revealed that this value did not differ significantly from 50%,  $t(59) = -1.29$ ,  $p = .204$ ,  $d = 0.17$ , indicating that participants understood the statistical constraint of the "Two Door" design.

**Signal detection analyses.** We next analyzed participants' judgments using signal detection analyses. As expected, participants displayed a significant response bias toward straight categorizations, judging only 39% of the targets to be gay ( $M_c = 0.32$ ,

$SD_c = 0.40$ ),  $t(59) = 6.17$ ,  $p < .001$ ,  $d = 0.80$ . We also noted that the magnitude of response bias differed across the backward and forward presentation orders ( $M_s = .21$  and  $.41$ , respectively),  $t(50.77) = -2.01$ ,  $p = .049$ ,  $d = 0.56$  ( $df$  corrected for unequal variance). Thus, we conducted additional analyses to test the magnitude of bias separately for each order. The bias toward straight categorizations remained significant in both the backward condition,  $t(27) = 4.09$ ,  $p < .001$ ,  $d = 0.78$ , and the forward condition,  $t(31) = 4.98$ ,  $p < .001$ ,  $d = 0.89$ .

Study 1 provided a stringent test of the straight categorization bias using a unique design in which participants believed each trial had equal chances of revealing a gay or straight face. Despite understanding the statistical constraints of this design, participants used the gay category label only 39% of the time, resulting in a notable response bias toward straight categorizations. It bears repeating that the magnitude of the effect differed somewhat as a function of presentation order: Although response bias was evident for both orders, it was noticeably stronger in the backward condition relative to the forward condition. This finding suggests that sexual orientation judgments depend in part on the targets perceivers encountered previously. We return to this possibility in Studies 5–6, but for now reiterate that Study 1 documented a significant bias toward straight judgments even when the design specified statistically independent trials with equal chances of revealing a gay or straight target.

## Study 2

Study 1 demonstrated that base rate matching does not fully account for the straight categorization bias, given that participants continued to exhibit a response bias when they believed each trial had equal chances of revealing a gay or straight target. It bears noting, however, that the Two Door design differed considerably from the methods used in prior research on social categorization. To corroborate our findings using methods more analogous to existing work, we conducted a second study in which participants received explicit base rate information before categorizing the sexual orientation of multiple targets in rapid succession. If the straight categorization bias is driven by perceivers' attempts to match base rates, we should see no evidence of bias when participants receive explicit base rate information before the task.

## Method

**Participants.** There were 126 Internet users who completed the study ( $M_{age} = 34.21$ ; 45% women; 75% White; 87% straight).

**Stimuli.** Stimuli were identical to Study 1.

**Procedure.** We recruited Mechanical Turk users for a study about their impressions of other people, with no mention of base rates or sexual orientation. After providing consent, participants were redirected to the survey hosting website Qualtrics, where they were randomly assigned to one of four conditions: 25% base rate ( $n = 32$ ), 50% base rate ( $n = 28$ ), 75% base rate ( $n = 36$ ), or control (unspecified base rate;  $n = 30$ ). All participants learned we were interested in their perceptions of others' sexual orientations. In the conditions specifying base rates, participants received additional information about the number of gay targets in the stimulus set. For example, participants in the 25% condition were told: "You should know that one quarter (25%), or 10 of the 40 people



that you will see, identify as gay in real life. Try to keep this in mind as you make your judgments.” The photographs remained identical across conditions, enabling a direct comparison of response bias across stimulus sets with identical targets but different base rate information.

Stimuli were presented individually in random order. As before, participants were given unlimited time to categorize each face and they received no feedback about the accuracy of their judgments. After completing all judgments, participants responded to a manipulation check: “Based on the information provided at the beginning of the study, what percentage of the people you saw identify as gay in real life?” Finally, participants reported demographic information before being debriefed.

## Results and Discussion

Study 2 aimed to provide a second test of the possibility that the straight categorization bias reflects participants’ attempt to calibrate their judgments to an expected base rate. If this hypothesis is correct, then explicitly informing participants of the base rate of gay targets in a stimulus set should eradicate the straight categorization bias. We used signal detection analyses to compare the magnitude of the response bias against the expected value given the base rate specified within each condition.

**Manipulation check.** Overall, 71% of participants correctly identified the base rate of gay targets based upon the information provided. This percentage was somewhat lower than we expected. Upon further inspection, it appears that incorrect responses were driven primarily by participants in the control condition who guessed at the base rate rather than selecting the “not specified” option. Discarding control participants who failed the manipulation check, 83% of the remaining participants provided correct responses. Nevertheless, the pattern of results replicated when excluding participants who failed the manipulation check, so we retained all responses.

**Signal detection analyses.** Participants displayed a significant response bias toward straight categorizations overall ( $M_c = 0.64$ ,  $SD_c = 1.85$ ),  $t(125) = 3.85$ ,  $p < .001$ ,  $d = 0.35$ . Because participants received different base rate information, however, we should expect response bias to vary across conditions. Indeed, analyzing  $c$  values using a one-way analysis of variance (ANOVA) with Base Rate Condition as a between-subjects factor indicated that response bias varied across base rate conditions,  $F(3, 122) = 4.88$ ,  $p = .003$ ,  $\eta_p^2 = 0.11$  (see Figure 1). We explored this variability further by conducting one-sample  $t$  tests within each condition to determine whether the observed response bias differed from the expected value given the specified base rate. In the 25% condition, the mean response bias was not significantly different from the expected value of 1.16 ( $M_c = 1.23$ ,  $SD_c = 2.29$ ),  $t(31) = 0.18$ ,  $p = .859$ ,  $d = 0.03$ . In the 50% condition, the mean response bias was significantly higher than the expected value of 0 ( $M_c = 0.34$ ,  $SD_c = 0.84$ ),  $t(27) = 2.18$ ,  $p = .039$ ,  $d = 0.40$ . In the 75% condition, the mean response bias was significantly higher than the expected value of  $-1.16$  ( $M_c = -0.15$ ,  $SD_c = 1.59$ ),  $t(35) = 3.82$ ,  $p = .001$ ,  $d = 0.64$ . This pattern indicates that response bias decreased as the presumed base rate increased. As a direct test of this effect, we regressed Response Bias onto Base Rate Condition (1 = 25%, 2 = 50%, and 3 = 75%). Indeed, response bias

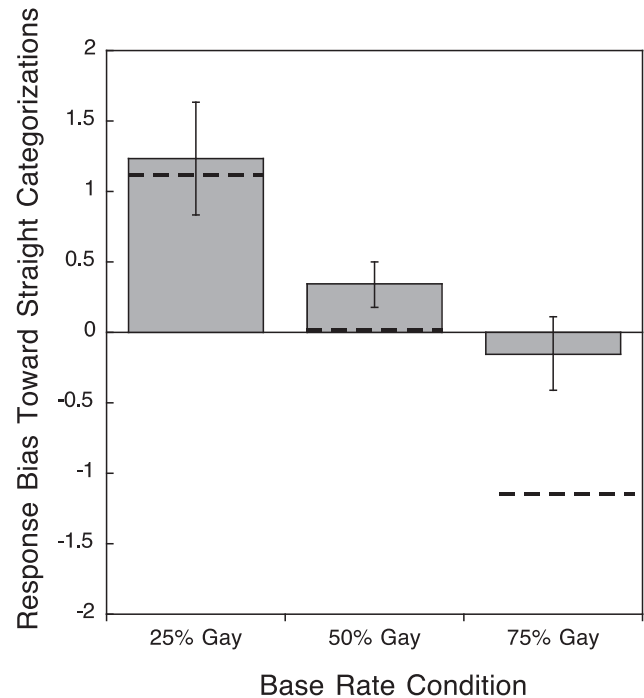


Figure 1. Effect of base rate knowledge on response bias in Study 2. Dashed lines indicate the expected level of response bias if perceivers were to match categorizations to the specified base rate. Participants in the 25% Base Rate Condition did not show a significant response bias, but participants in the 50% and the 75% Base Rate Condition showed a significant bias toward straight categorizations relative to the base rate information they received.

decreased in a linear fashion as the presumed base rate of gay faces increased,  $B = -0.67$ ,  $SE = 0.21$ ,  $t = -3.34$ ,  $p = .001$ .

We conducted additional analyses to explore response tendencies in the control condition for which the base rate was unspecified. Here, participants showed a significant response bias toward straight categorizations ( $M_c = 1.21$ ,  $SD_c = 1.97$ ),  $t(29) = 3.37$ ,  $p = .002$ ,  $d = 0.61$ . It is noteworthy that the mean response bias in the unspecified condition was nearly identical to the mean response bias in the 25% base rate condition ( $M_s = 1.21$  and  $1.23$ , respectively). Further analyses of participants’ responses to the manipulation check provide additional clarity for this effect. Recall that many of the participants in the unspecified condition failed the manipulation check because they guessed at the base rate rather than selecting the “not specified” option. Of those incorrect guesses, nearly half (47%) indicated a base rate of 25%. This finding suggests that, when lacking clear base rate information, a common assumption is that one-quarter of the targets identify as gay.

Study 2 revealed that base rate matching exerts some impact on the outcomes of sexual orientation categorization. Indeed, the magnitude of response bias varied across participants who received different base rate information about an identical stimulus set, with response bias decreasing in a linear fashion as the stated base rate increased from 25 to 75%. Furthermore, participants did not show a response bias when they believed the base rate to be 25%. This result is particularly compelling alongside the observa-

tion that participants in the unspecified condition tended to guess the base rate was 25%. Participants were especially prone to show a response bias toward straight categorizations when the empirical base rate exceeded their baseline assumption of 25% gay targets.

While participants partially calibrated their sexual orientation judgments to the explicitly provided base rate, they failed to adjust their categorizations sufficiently. In all but the 25% base rate condition, participants showed a strong and significant bias toward straight categorizations even when they were able to correctly indicate the stated base rate in a manipulation check. Therefore, while base rate matching appears to inform sexual orientation judgments to some extent, it does not provide a complete explanation for the straight categorization bias.

### Study 3

Because the stimulus set was fixed in Study 2, the base rate information participants received was technically invalid in all but the 50% condition. Manipulating presumed base rates in this manner provided control for testing our hypotheses about base rate matching as an explanation for the straight categorization bias, but it also created an unrealistic scenario that might have conflicted with participants' observation of diagnostic cues displayed by targets. Furthermore, our prior studies were unique in that they provided participants with clear base rate information about the stimulus set. While this approach allowed us to test whether participants match their categorization tendencies to explicitly stated base rates, it precluded a test of participants' sensitivity to base rates in the more common scenario when they are left unspecified. Study 3 tested whether base rate matching accounts for the straight categorization bias by altering the actual base rates of gay targets in the stimulus set without informing participants of this variability.

### Method

**Participants.** There were 203 Internet users who completed the study ( $M_{age} = 34.74$ ; 50% women; 76% White; 89% straight).

**Stimuli.** Stimuli consisted primarily of those described in Studies 1 and 2, but the manipulation of the actual base rate of gay targets in the stimulus set required the inclusion of several additional faces from Freeman et al. (2010). In the 25% base rate condition, we replaced 10 gay faces with 10 new straight faces (5 men, 5 women). In the 75% base rate condition, we replaced 10 straight faces with 10 new gay faces (5 men, 5 women). As before, all photographs were standardized to  $200 \times 288$  pixels and depicted White individuals devoid of facial hair, jewelry, and visible tattoos.

**Procedure.** We recruited Mechanical Turk users for a study about their impressions of other people, with no mention of base rates or sexual orientation. After providing consent, participants were redirected to the survey hosting website Qualtrics, where they were randomly assigned to one of three conditions: 25% base rate ( $n = 68$ ), 50% base rate ( $n = 67$ ), or 75% base rate ( $n = 68$ ). Across conditions, participants learned we were interested in their perceptions of others' sexual orientations, but they received no explicit information about base rates. Instead, base rate information was provided implicitly based upon participants' experience of the stimulus set as they made judgments. After completing their

judgments, participants were asked to indicate what percentage of the targets they believed were actually gay. All other procedures were identical to Study 2.

### Results and Discussion

Study 3 tested whether participants adjust their categorization tendencies to match the actual base rate of gay targets in a stimulus set. If perceivers infer base rates from visual exposure to diagnostic cues, then the straight categorization bias should become less pronounced as the number of gay targets in the stimulus set increases. If not, then the straight categorization bias should persist across conditions. We tested these possibilities by comparing the magnitude of response bias to the expected value given the base rate of gay targets in each condition.

Perceivers displayed a response bias toward straight categorizations in the sample overall ( $M_c = 0.63$ ,  $SD_c = 0.54$ ),  $t(202) = 16.76$ ,  $p < .001$ ,  $d = 1.17$ . Again, however, analyzing  $c$  values using a one-way ANOVA with Base Rate Condition as a between-subjects factor revealed that response bias varied across base rate conditions,  $F(2, 200) = 2.69$ ,  $p = .071$ ,  $\eta_p^2 = 0.03$  (see Figure 2). We explored this variability by conducting one-sample  $t$  tests within each condition to determine if the observed response bias differed from the expected value given the base rate. In the 25% condition, the mean response bias was lower than the expected value of 1.16 ( $M_c = 0.52$ ,  $SD_c = 0.56$ ),  $t(68) = -9.45$ ,  $p < .001$ ,  $d = 1.14$ . In the 50% condition, the mean response bias was higher than the expected value of 0 ( $M_c = 0.66$ ,  $SD_c = 0.52$ ),  $t(65) =$

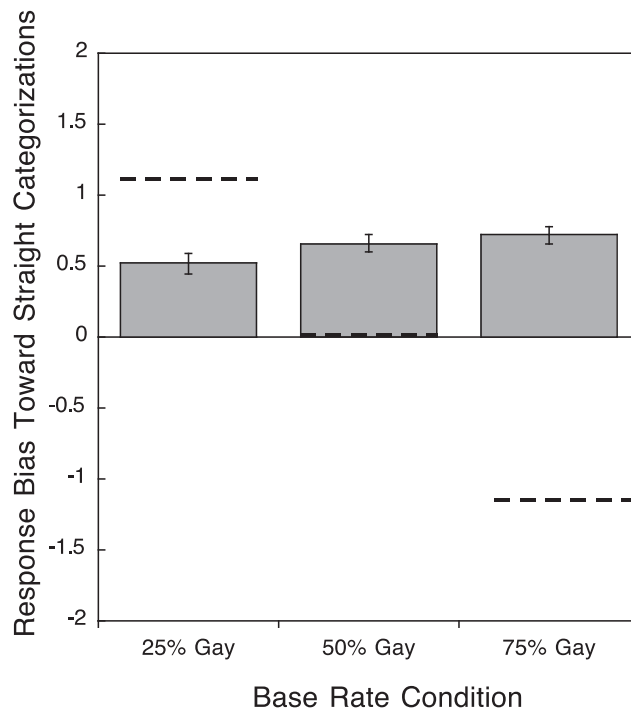


Figure 2. Effect of stimulus base rates on response bias in Study 3. Dashed lines indicate the expected level of response bias given the true base rate of gay faces in the stimulus set. Participants in the 50% and the 75% Base Rate Condition showed a significant bias toward straight categorizations relative to the base rate information they received.

10.44,  $p < .001$ ,  $d = 1.27$ . In the 75% condition, the mean response bias was again higher than the expected value of  $-1.16$  ( $M_c = 0.72$ ,  $SD_c = 0.52$ ),  $t(67) = 27.39$ ,  $p < .001$ ,  $d = 3.62$ .

The findings described above indicate that the straight categorization bias *increased* as the number of gay targets in the sample increased. As a direct test of this effect, we regressed Response Bias onto Base Rate Condition (1 = 25%, 2 = 50%, and 3 = 75%). Indeed, response bias increased in a linear fashion as the actual base rate of gay faces increased,  $B = 0.10$ ,  $SE = 0.05$ ,  $t = 2.26$ ,  $p = .025$ . One possible explanation for this finding is the fact that perceivers received no explicit information about the empirical base rate at the beginning of this study; thus, they could not calibrate their judgments to the base rate until they had accumulated some experience with the stimulus set. To test this possibility, we recomputed signal detection analyses for the first 20 stimuli and the last 20 stimuli each participant encountered. The pattern for the first 20 stimulus exposures was identical to the overall pattern described above. Specifically, analyzing  $c$  values using a one-way ANOVA with Base Rate Condition as a between-subjects factor revealed that response bias varied across base rate conditions,  $F(2, 200) = 3.39$ ,  $p = .036$ ,  $\eta_p^2 = 0.03$ . In the 25% condition, the mean response bias was lower than the expected value of 1.16 ( $M_c = 0.48$ ,  $SD_c = 0.57$ ),  $t(68) = -10.00$ ,  $p < .001$ ,  $d = 1.19$ . In the 50% condition, the mean response bias was higher than the expected value of 0 ( $M_c = 0.58$ ,  $SD_c = 0.47$ ),  $t(65) = 10.04$ ,  $p < .001$ ,  $d = 1.23$ . In the 75% condition, the mean response bias was again higher than the expected value of  $-1.16$  ( $M_c = 0.71$ ,  $SD_c = 0.48$ ),  $t(67) = 32.02$ ,  $p < .001$ ,  $d = 3.89$ . Regression analysis revealed that response bias increased in a linear fashion as the actual base rate of gay faces increased,  $B = 0.11$ ,  $SE = 0.04$ ,  $t = 2.61$ ,  $p = .010$ .

We conducted an identical series of analyses for the last 20 stimulus exposures. Again, response bias varied somewhat across base rate conditions  $F(2, 200) = 3.43$ ,  $p = .034$ ,  $\eta_p^2 = 0.03$ . In the 25% condition, the mean response bias was lower than the expected value of 1.16 ( $M_c = 0.52$ ,  $SD_c = 0.71$ ),  $t(68) = -7.46$ ,  $p < .001$ ,  $d = 0.90$ . In the 50% condition, the mean response bias was higher than the expected value of 0 ( $M_c = 0.69$ ,  $SD_c = 0.51$ ),  $t(65) = 11.00$ ,  $p < .001$ ,  $d = 1.35$ . In the 75% condition, the mean response bias was again higher than the expected value of  $-1.16$  ( $M_c = 0.78$ ,  $SD_c = 0.55$ ),  $t(67) = 29.03$ ,  $p < .001$ ,  $d = 3.53$ . As before, regression analysis revealed that response bias increased in a linear fashion as the actual base rate of gay faces increased,  $B = 0.13$ ,  $SE = 0.05$ ,  $t = 2.59$ ,  $p = .010$ .

Finally, as a direct comparison of response bias across the two halves of the study, we subtracted each participant's response bias for the first 20 stimuli they judged from their response bias for the second 20 stimuli they judged. We then conducted a one-sample  $t$  test on this difference score, which revealed that the straight categorization bias was slightly stronger for the second 20 faces relative to the first 20 faces ( $M = 0.07$ ,  $SD = 0.54$ ),  $t(202) = 1.95$ ,  $p = .053$ ,  $d = 0.13$ . This slight increase in response bias was most pronounced for participants in the 50% base rate condition ( $M = 0.11$ ,  $SD = 0.41$ ),  $t(65) = 2.14$ ,  $p = .036$ ,  $d = 0.27$ , but not significant for participants in the 25% base rate condition ( $M = 0.04$ ,  $SD = 0.65$ ),  $t(68) = 0.50$ ,  $p = .622$ ,  $d = 0.06$ , or the 75% base rate condition ( $M = 0.08$ ,  $SD = 0.55$ ),  $t(67) = 1.18$ ,  $p = .242$ ,  $d = 0.15$ .

The above analyses cast doubt on the hypothesis that participants extract base rate information from a stimulus set and use that information to guide their judgments. If participants were attuned to base rate information gleaned from their cumulative exposure to a stimulus set, response bias should have decreased as the number of gay faces increased, and this effect should have been more pronounced later in the study as opposed to earlier. We observed neither of those patterns here. An alternative hypothesis is that, in the absence of explicit information about the stimulus set, participants relied on their own intuitions about base rates given that the judgment task required them to sort faces into one of two categories. In this situation, participants may have assumed a base rate of 50% and tailored their categorizations accordingly, resulting in an increasingly strong response bias as the true base rate of gay targets increased. To test this possibility, we examined participants' presumed base rates given their estimate of the number of targets they believed were actually gay. A majority of participants (71%) indicated a base rate of 50%, and these estimates did not vary as a function of the true number of gay targets in the stimulus set,  $F(2, 200) = 0.51$ ,  $p = .601$ ,  $\eta_p^2 = 0.01$ . Even more informative is the fact that presumed base rates were strongly tethered to response tendencies. As participants believed the number of gay targets in the stimulus set increased, the straight categorization bias decreased,  $r(203) = -0.45$ ,  $p < .001$ . These findings suggest that *presumed* base rates guide sexual orientation judgments when explicit base rate information is not available. Still, while participants' judgments were correlated with their presumed base rate of gay targets, they nevertheless showed a straight categorization bias relative to their own beliefs: There was a significant response bias toward straight categorizations among participants who presumed a 50% base rate of gay targets ( $M = 0.65$ ,  $SD = 0.45$ ),  $t(144) = 17.50$ ,  $p < .001$ ,  $d = 1.44$ .

Study 3 provided additional evidence of a straight categorization bias in sexual orientation categorization. Participants did not accurately calibrate their categorizations to the true base rate of gay targets in the stimulus set; instead, they tended to infer a 50% base rate and make judgments proportional to this inference rather than to the stimuli they encountered. It also bears repeating that the straight categorization bias still emerged when we tested response bias against this presumed base rate. Participants underutilized the gay category label relative to their own beliefs as well as the true number of gay targets in each stimulus set.

In combination, Studies 1–3 revealed a significant bias toward straight categorizations that replicated across several distinct judgment paradigms. The data indicated that base rate matching partially contributes to this bias, insofar as participants calibrated their judgments to base rate information that was either explicitly provided or inferred from the stimulus set. However, base rate matching provided an incomplete explanation for the bias. Perceivers continued to show a response bias toward straight categorizations when they believed each trial had a 50% chance of revealing a gay target (Study 1), when they received and remembered explicit base rate information about the stimulus set in question (Study 2), and when they judged targets from stimulus sets with varying base rates (Study 3). Participants even underutilized the gay category label relative to their own explicitly stated beliefs about the base rate of gay targets in the stimulus set (Study 3). Thus, while base rate matching certainly impacts sexual

orientation judgments, it does not provide a complete explanation for the response bias observed here.

### Study 4

Our initial studies revealed a response bias toward straight judgments in sexual orientation categorization that is not fully accounted for by base rate matching. Because base rates do not provide a full explanation for the bias, other social-cognitive factors must also contribute to it. The remaining studies test a second account based upon prior insights about the perceptual mechanisms guiding sexual orientation judgments. In particular, numerous studies have shown that perceivers rely on a gender inversion heuristic to make sexual orientation judgments, categorizing targets whose features adhere to those expected for their sex (masculine men, feminine women) as straight and targets whose features depart from those expected for their sex (feminine men, masculine women) as gay (Freeman et al., 2010; Johnson et al., 2007; Lick et al., 2013). Although use of this gender inversion heuristic appears robust, it is important to note that the mapping of gendered features onto sexual orientations is imperfect—not all hyper-gendered targets are straight, and not all androgynous targets are gay. Rather, relying on gendered features to categorize sexual orientation could be construed as a form of stereotyping wherein perceivers generalize beliefs about the gendered features of gay men and lesbians in the population to individual exemplars (Cox, Devine, Bischmann, & Hyde, 2015).

While recent studies have provided important insights about the heuristics guiding sexual orientation judgments, researchers have yet to pinpoint the precise threshold of gendered features compelling gay categorizations. Two possibilities seem reasonable. First, as depicted in Figure 3A, perceivers may evenly map gendered features onto sexual orientation judgments. Given equivalent base rates for gay and straight targets, this scenario implies that individuals whose features are androgynous relative to the rest of the sample will be categorized as gay whereas individuals whose features are hyper-gendered relative to the rest of the sample will be categorized as straight. A second possibility, depicted in Figure 3B, is that perceivers possess an uneven mapping of gendered features onto sexual orientation judgments. In this scenario, the threshold for gay categorizations falls well below the midpoint of the range of gendered features for the sample, resulting in a relatively small range of cues supporting gay categorizations. This limited range would yield a response bias because few targets exhibit features that surpass perceivers' threshold for categorizing others as gay. In the context of a study in which the base rate of gay targets is specified at 50%, the top figure represents an idealized mapping of gendered features onto sexual orientation categorizations. Indeed, because norms for gendered facial features are tethered to the specific set of faces encountered within a task (Lick & Johnson, 2014b), participants should calibrate their judgments accordingly: Targets whose features fall above the midpoint of gendered features for the sample should be categorized as straight, and targets whose features fall below the midpoint of gendered features for the sample should be categorized as gay. Study 4 tested this assumption about the heuristic underlying sexual orientation judgments.

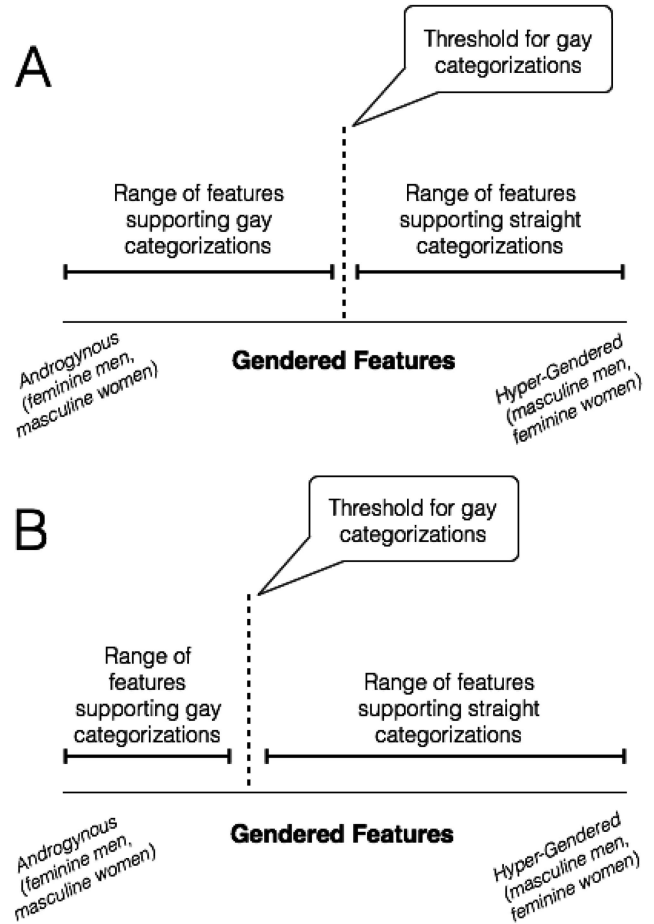


Figure 3. Conceptual diagrams representing two possibilities for the range of gendered cues perceivers use to categorize targets as lesbian/gay and straight. The top diagram (A) depicts a scenario in which perceivers have a balanced mapping of gendered features onto sexual orientation categorizations; targets with relatively androgynous features (below the midpoint of the gender spectrum) are considered gay. The bottom diagram (B) depicts a scenario in which perceivers have an uneven mapping of gendered features onto sexual orientation categorizations; there is a limited range of features compelling gay categorizations.

### Method

**Participants.** There were 271 Internet users who completed the study ( $M_{age} = 35.86$ ; 47% women; 75% White; 91% straight).

**Stimuli.** We created stimuli that varied systematically in gendered appearance using a morphing paradigm. We began by creating an androgynous composite face with equally male-typed and female-typed features. Specifically, we selected front-facing photographs of 18 White men and 18 White women with neutral expressions from the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015). We then used Abrosoft morphing software to place 112 landmark points around each face before mathematically averaging the images into a single composite face. By averaging 36 faces into a single composite face, we reduced the likelihood of idiosyncratic features from a single face exerting undue influence on the morphs. After creating the androgynous composite face, we



selected additional front-facing photographs of 5 White men and 5 White women with neutral expressions from the same database. We used these images as parent faces to create 10 continua that varied systematically in their gendered appearance. Specifically, we placed 112 landmark points around each parent face and then morphed them toward and away from the androgynous composite in 11 evenly spaced intervals to exaggerate the physical differences between the male/female parent faces and the androgynous composite face. This procedure yielded 110 faces (55 men, 55 women) that varied systematically in their gendered appearance from  $-100\%$  identity strength (androgynous) to  $100\%$  identity strength (hyper-gendered). The resulting images were oval-cropped to remove hair cues and resized to  $199 \times 281$  pixels (see Figure 4).

**Procedure.** We recruited Mechanical Turk users for a study about their perceptions of other people, with no mention of gender or sexual orientation. After providing consent, participants viewed either male or female test faces in random order and judged the sexual orientation of each one twice to ensure reliability.<sup>3</sup> We informed participants that some of the faces would look similar to one another, but asked them to judge each one individually, regardless of their responses to previous faces. To avoid potential confounds related to participants' attempts to match population base rates, we explicitly instructed them that half of the targets were gay and asked them to keep this in mind throughout the study. Upon completing their judgments, participants answered a manipulation check testing their knowledge of the base rate and reported demographic information before being debriefed.

## Results and Discussion

Our primary aim in Study 4 was to estimate the range of gendered cues perceivers use to categorize others as gay versus straight. We did so by fitting psychometric curves to the judgment data and locating the point of subjective equality (PSE)—that is, the point on the gender morph continuum where gay and straight category judgments were equally likely. To do so, we coded the test stimuli according to an 11-point scale indicating the degree of gender morph ( $-100$  [androgynous],  $-80$ ,  $-60$ ,  $-40$ ,  $-20$ ,  $0$  [parent face],  $20$ ,  $40$ ,  $60$ ,  $80$ ,  $100$  [hyper-gendered]). We then aggregated each participant's sexual orientation judgments across target identities to form the proportion of targets they categorized as gay at each level of gender morph. Next, we used Prism software to fit psychometric functions to the aggregated data on the basis of a cumulative Gaussian distribution, which allowed us to calculate PSE. If the average PSE falls above the scale midpoint, then there is a limited range of features compelling gay categorizations.

Fifty-nine participants failed a manipulation check testing their knowledge of the base rate information supplied at the beginning of the study, and we excluded their responses to ensure that findings could not be attributed to participants' attempt to match the low population base rate of sexual minorities. Moreover, 162 participants had nonconvergent or very poorly fitting psychometric curves ( $R^2 < .60$ ) that we excluded from PSE analysis. This may seem like a large number of exclusions, but it was necessary for the analysis of psychometric response data. Indeed, we could not analyze responses from participants with nonconvergent curves because we could not calculate a PSE, and it would have been

inappropriate to include the curves with severe misfit because such misfit suggests that participants were not discriminating between the response options *lesbian/gay* and *straight* with regard to gendered features. Knowing that judgments of concealable identities are prone to misfitting curves, we intentionally recruited enough participants to ensure sufficient power after exclusions. After removing participants who failed the manipulation check and those with poorly fitting curves, we still had an effective sample of 94 participants. A power analysis revealed that 67 participants were required to achieve 80% power to detect small-to-medium effects ( $d = 0.35$ ; based on pilot test) in a one-sample  $t$  test when  $\alpha = .05$ . Therefore, the study was appropriately powered even after necessary data exclusions. That said, the pattern of results replicates if we include participants who failed the manipulation check as well as in secondary analyses that included those with misfitting curves (see below).

We began by testing whether the PSE differed as a function of target sex. The average PSE did not differ between male targets ( $M = -8.40$ ,  $SD = 24.07$ ) and female targets ( $M = -9.48$ ,  $SD = 28.44$ ),  $t(92) = 0.20$ ,  $p = .844$ ,  $d = 0.04$ , so we report analyses for both sexes combined. A one-sample  $t$  test indicated that the PSE fell significantly below the midpoint of the gender morph continuum ( $M = -8.80$ ,  $SD = 25.64$ ),  $t(93) = -3.33$ ,  $p = .001$ ,  $d = 0.34$  (see Figure 5). Participants required a face that was somewhat extreme in gender inversion before rendering a gay judgment.

Removing participants with misfitting psychometric curves might raise concerns that we limited our sample to participants who relied heavily on gendered cues when categorizing sexual orientation, increasing our chances of finding evidence for a limited range of gendered cues supporting gay categorizations. We therefore conducted an additional set of analyses that allowed us to include data from all participants who passed the manipulation check. We restructured participants' responses into long format so that each trial occupied one line of the dataset. Then, using a generalized estimating equation (GEE; Zeger & Liang, 1986) to account for the nested data, we regressed Sexual Orientation Categorization ( $0 = straight$ ,  $1 = gay$ ) onto Gender Morph ( $-100 = androgynous$  to  $100 = hyper-gendered$ ). As expected, the slope for Gender Morph was negative and significant,  $B = -0.01$ ,  $SE < 0.01$ ,  $z = -4.61$ ,  $p < .001$ , indicating that participants applied a gender inversion heuristic whereby gay categorizations became more likely as faces became more androgynous. More important for our purposes is the intercept, which indicates the logit value for gay categorizations when Gender Morph is equal to 0. The intercept was significantly below zero,  $B = -0.09$ ,  $SE = 0.03$ ,  $z = -2.98$ ,  $p = .003$ , indicating that participants categorized targets at the midpoint of the gender scale as straight even when the 50% base rate should have rendered these categorizations gay.

Study 4 revealed that the range of gendered features for which participants made gay categorizations was smaller than the range of gendered facial features for which they made straight categori-

<sup>3</sup> Participants evaluated either male or female faces because categorizing both sexes in a single study may have caused confusion about the sex of the androgynous face. Specifying target sex a priori served to clarify that the androgynous faces were gender-atypical members of a given sex.

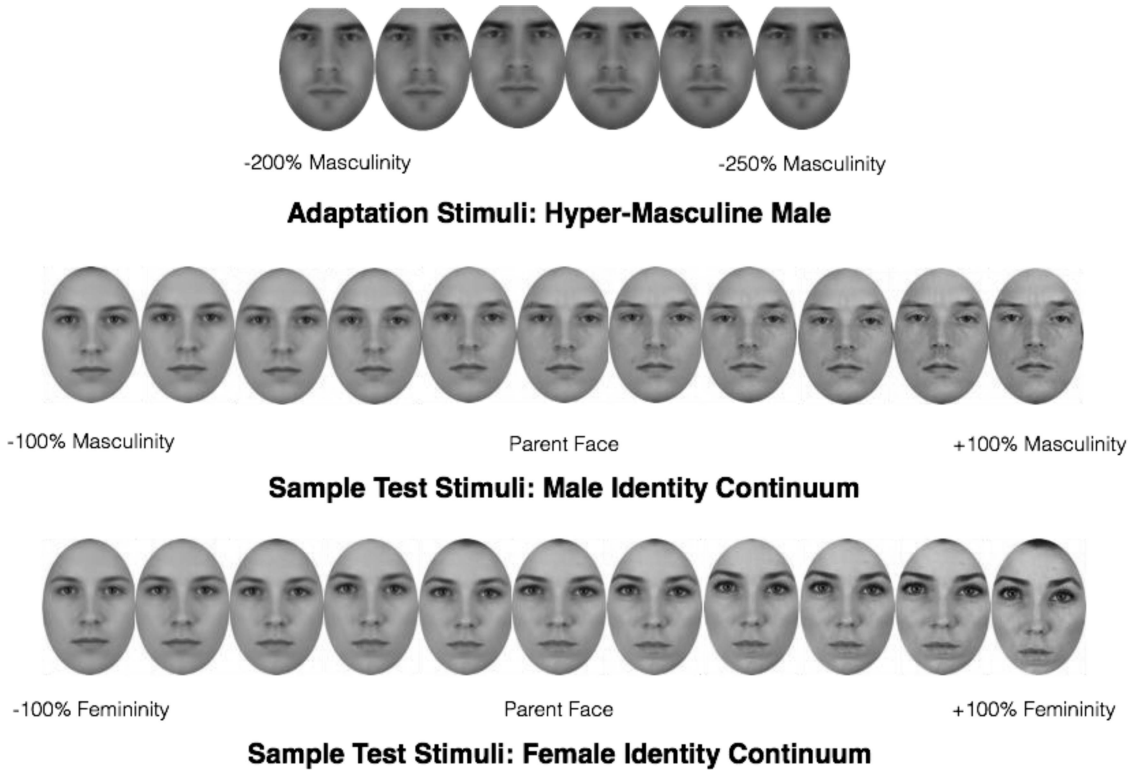


Figure 4. Sample stimuli from Studies 4–5. Adaptation stimuli were hyper-masculine male faces from Zhao et al. (2011). Test stimuli were male and female face continua created by morphing parent faces from Ma et al. (2015) toward and away from an androgynous composite comprised of equal numbers of male and female faces.

zations, despite being given equal base rate expectations for straight and gay targets. If perceivers relied on a balanced decision criterion when rendering sexual orientation judgments, they should have calibrated their gender heuristic to the features of the stimulus set, categorizing the targets whose features were more gender adherent than the midpoint as straight and the targets whose features were less gender adherent than the midpoint as gay. Instead, participants tended to require more extreme deviations to compel gay categorizations relative to straight categorizations. Although the magnitude of the effect was modest, it was highly reliable, providing evidence for a limited range of cues supporting gay categorizations. As such, this discovery sheds new light on the gendered heuristic guiding sexual orientation categorization while simultaneously documenting perceptual underpinnings of the straight categorization bias. Fewer people possess features that match perceivers' expectations of sexual minority individuals than possess features that match perceivers' expectations of straight individuals, resulting in an upwardly biased threshold for gay categorizations.

Two considerations about the logic underlying our conclusions deserve mention. The first is that it might be inappropriate to expect perceivers to set the threshold for sexual orientation categorization at the midpoint of a range of gendered features. If the population of gay men and lesbians is rather small, one might argue that perceivers *should* place their threshold for gay categorizations left of center. However, this argument is faulty because (a) the instructions for Study 4 clearly specified a base rate of 50%

gay targets, and (b) we excluded participants who failed a manipulation check testing their knowledge of this constraint. Thus, the results cannot readily be attributed to the fact that sexual minorities make up a relatively small portion of the population. The second consideration is that it might be unrealistic to expect perceivers to apply sexual orientation categorizations evenly across the range of gendered features included in this study, given previous work indicating that perceivers generally expect lesbian and gay targets to appear somewhat androgynous. Because the center of the gender distribution consisted of the original parent faces, most of whom presumably identified as straight, it might seem as if the point of subjective equality should be shifted left. Again, however, our analyses pertain only to participants who correctly recalled the 50% base rate constraint provided at the beginning of the study. Moreover, the typicality of gendered features is relative to the context of faces in which an exemplar is viewed (Lick & Johnson, 2014b). Because perceivers were informed the base rate of gay targets was 50%, they should have calibrated their judgments to utilize the observable gendered features accordingly. That is, the gender inversion heuristic should have led perceivers to place their perceptual threshold for gay categorizations at a location that would split the range of gendered features into even halves. However, even when participants acknowledged the 50% base rate, their threshold for gay categorizations fell below the midpoint of the gender continuum, indicating a limited range of cues supporting gay categorizations.

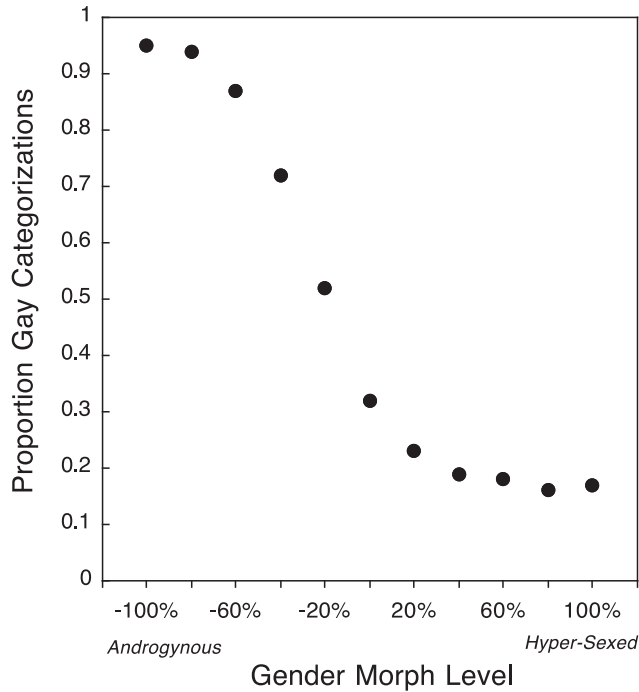


Figure 5. Overall psychometric function displaying the proportion of gay categorizations as a function of stimulus gender morph in Study 4.

### Study 5

Study 4 revealed that a limited range of gender cues supports gay categorizations even when instructions specified equal numbers of gay and straight targets. While these insights provide novel information about a mechanism by which sexual orientation categorizations become biased, it remains unclear why such a limited range exists. One possibility is that perceivers have relatively limited exposure to sexual minority individuals, with many prominent exemplars in film and television being highly stereotyped (Dyer, 1999; Hart, 2000; Raley & Lucas, 2006). Repeated exposure to such stereotypical representations may lead perceivers to believe that gay men are androgynous, which could shift the threshold for gay categorizations toward the androgynous end of the gender continuum and extend the perceptual range beyond normal anthropometric variation. That is, perceivers may calibrate the features they consider to be consistent with the category gay based upon their experiences with known members of the category, many of whom exhibit exaggerated gender stereotypes.

If the threshold for assigning gay category labels is indeed based on perceivers' prior experiences, then it should shift after experimental exposure to targets with features stereotyped in the opposite direction. Indeed, research in vision science has revealed that perceivers' norms for face categories are malleable, such that exposure to a set of features (*adaptation*) recalibrates face space and alters the perception of subsequent targets (*aftereffects*; Webster, 2012). In particular, adapting perceivers to faces with hyper-gendered features shifts the perceived norm for faces toward a hyper-gendered exemplar, causing previously neutral faces to appear altered in the opposite direction of adaptation (Lick & John-

son, 2014b). Applied to the current work, these findings suggest that exposing perceivers to hyper-gendered faces may cause previously neutral faces to appear more androgynous and therefore shift the perceptual threshold for gay categorizations and eliminate the limited range of gendered cues used to support gay categorizations documented previously.

### Method

**Participants.** There were 117 Internet users who completed the study ( $M_{age} = 35.76$ ; 45% women; 75% White; 94% straight).

**Stimuli.** We only included male faces for the sake of efficiency because effects did not vary as a function of target sex in Study 4 or in prior work on visual adaptation to gendered facial features. Test stimuli were the male face continua described in Study 4. Adaptation stimuli were six hyper-masculine male faces from prior research (Zhao, Seriès, Hancock, & Bednar, 2011; Figure 4), which were created by averaging the faces of 150 young adult White men and exaggerating their gendered features away from a female composite up to 250%. Adaptation images were cropped to remove hair cues and presented at a smaller size than test images ( $149 \times 193$  pixels) to ensure that aftereffects reflected high-level changes in gender representation as opposed to low-level changes in retinotopic processing.

**Procedure.** We recruited Mechanical Turk users for a study about their perceptions of other people, with no mention of gender or sexual orientation. At pretest, participants viewed the test faces in random order and categorized each one's sexual orientation. As before, we instructed participants that half of the faces identified as gay and asked them to keep this in mind while making judgments. Next, participants underwent a visual adaptation in which six hyper-masculine male faces were repeatedly displayed in random order for 3 s each for a total of 3 min. We opted only to adapt perceivers to hyper-masculine faces because prior work on visual adaptation revealed strong and consistent gender aftereffects after adaptation to both masculine and feminine faces (Lick & Johnson, 2014b; Webster, Kaping, Mizokami, & Duhamel, 2004). Also consistent with prior research, one-quarter of the adaptation faces were marked with a faintly colored circle, which participants were instructed to identify by pressing spacebar as quickly as possible when they appeared. We did not record the speed with which participants identified marked faces; the detection task was a ruse intended to ensure visual engagement throughout the adaptation period. Finally, participants completed posttest judgments that were identical to pretest judgments with one addition: We included a brief top-up adaptation (randomly selected adapting face presented for 2,000 ms) before each test face to ensure that adaptation continues throughout the posttest period. After making their judgments, participants completed a manipulation check testing their knowledge of the specified base rate, reported demographic information, and were debriefed.

### Results and Discussion

Study 5 sought to determine whether perceptual exposure alters the perceptual threshold at which perceivers reliably categorize unknown others as gay. Specifically, we tested whether visual adaptation to hyper-masculine male faces shifted this threshold toward the center of the gender continuum, broadening the range



of cues supporting gay categorizations. We did so by fitting psychometric curves to the judgment data and calculating separate PSEs for pretest and posttest, using the same methods described in Study 4.

We excluded 14 participants who failed a manipulation check testing their knowledge that half of the targets identified as gay, 47 participants who had nonconvergent or very poorly fitting curves at pretest, and 41 participants who had nonconvergent or very poorly fitting curves at posttest ( $R^2 < .60$ ).<sup>4</sup> Again, this may seem like a large number of exclusions, but we intentionally recruited enough participants to ensure sufficient power (67 participants required to achieve 80% power to detect small-to-medium effects in a dependent-samples  $t$  test when  $\alpha = .05$ ; see above). That said, the pattern of results replicates if we include participants who failed the manipulation check as well as in secondary analyses that included those with misfitting curves (see below).

We first replicated our previous findings by analyzing the Pretest PSE. A one-sample  $t$  test revealed that the PSE was significantly below the midpoint of the gender continuum ( $M = -5.08$ ,  $SD = 18.33$ ),  $t(63) = -2.22$ ,  $p = .030$ ,  $d = 0.28$ . Thus, there was a smaller range of gendered cues compelling gay categorizations than straight categorizations at pretest. Next, we examined the PSE after adaptation. Unlike pretest, the Posttest PSE was not significantly different from the midpoint of the gender continuum ( $M = 2.27$ ,  $SD = 24.17$ ),  $t(67) = 0.78$ ,  $p = .441$ ,  $d = 0.09$ . The range of gendered cues compelling gay and straight categorizations was equivalent after adaptation to hyper-masculine faces.

The above findings indicate that visual adaptation shifted the threshold for gay categorizations toward the masculine end of the gender spectrum, broadening the range of facial features that reliably compelled gay categorizations. To systematically test whether the PSE changed as a function of visual adaptation, we subtracted Pretest PSE from Posttest PSE. We then subjected this difference score to a one-sample  $t$  test. As expected, the PSE was significantly shifted toward the masculine end of the gender spectrum after adaptation to hyper-masculine faces ( $M = 7.50$ ,  $SD = 26.67$ ),  $t(59) = 2.18$ ,  $p = .033$ ,  $d = 0.28$  (see Figure 6).

We conducted additional analyses to allay concerns about data exclusions in the PSE analysis. Specifically, we examined the probability of gay categorizations as a function of gender morph, allowing us to include data from all participants who passed the manipulation check. Using GEEs to account for nested data, we regressed Sexual Orientation Categorization ( $0 = \text{straight}$ ,  $1 = \text{gay}$ ) onto Gender Morph ( $-100 = \text{androgynous}$  to  $100 = \text{hyper-gendered}$ ). At pretest, the slope for Gender Morph was negative and significant,  $B = -0.01$ ,  $SE < 0.01$ ,  $z = -7.73$ ,  $p < .001$ , indicating that participants applied a gender inversion heuristic whereby gay categorizations became more likely as faces became more androgynous. More important for our purposes is the intercept, which indicates the logit value for gay categorizations when Gender Morph = 0. As expected, the intercept value was significantly below zero,  $B = -0.21$ ,  $SE = 0.06$ ,  $z = -3.81$ ,  $p < .001$ , suggesting that participants categorized targets at the midpoint of the gender scale as straight even when the base rate constraint should have rendered these categorizations gay. At posttest, the slope for Gender Morph was again negative and significant,  $B = -0.01$ ,  $SE < 0.01$ ,  $z = -7.44$ ,  $p < .001$ , indicating that participants still applied a gender inversion heuristic. Critically, however, the intercept value was no longer significant; if anything,

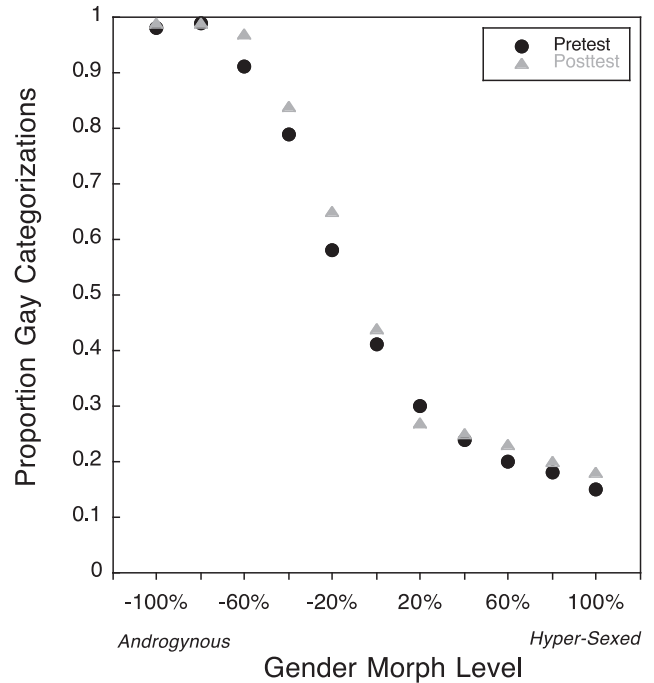


Figure 6. Overall psychometric function displaying the proportion of gay categorizations as a function of stimulus gender morph at pretest (black circles) and posttest (gray triangles) in Study 5. The point of subjective equality was shifted toward the center of the gender morph continuum after adaptation to hyper-masculine male faces.

it went in the opposite direction of the pretest intercept,  $B = 0.04$ ,  $SE = 0.07$ ,  $z = 0.56$ ,  $p = .578$ . These findings corroborate those from the PSE analysis, revealing that visual adaptation to hyper-masculine faces resulted in a more balanced mapping of gendered features to sexual orientation judgments.

Study 5 replicated our previous observation about the limited range of gendered cues supporting gay categorizations, and it extended that observation by demonstrating that the range is calibrated in part based upon perceptual experience. At pretest, participants displayed the same range restriction evident in Study 4, such that participants required a relatively androgynous appearance before categorizing others as gay. Critically, however, this threshold shifted as a function of visual exposure: Adapting perceivers to hyper-masculine male faces altered the point of subjective equality for sexual orientation categorizations, centering the perceptual threshold for gay category judgments. Of course, it is worth noting that participants in this study received a fairly strong dose of visual exposure—3 min of continuous adaptation to hyper-gendered faces. While this amount of adaptation was sufficient to shift the threshold for gay categorization, it remains unclear how quickly these effects buildup or decay, especially after a lifetime of exposure to stereotyped portrayals of gay men and lesbians in media (see Dyer, 1999; Hart, 2000; Raley & Lucas, 2006). Nevertheless, the findings serve as an existence proof of the phenomenon, revealing that the perceptual threshold for categorizing un-

<sup>4</sup> There was substantial overlap in the participants with misfitting curves at pretest and posttest.



known others as gay is based at least in part upon visual experience.

### Study 6

Studies 4 and 5 documented a limited range of gendered cues supporting gay categorizations and showed that this range is calibrated on the basis of recent experience. We have argued that this range restriction contributes to the straight categorization bias because there exist fewer targets whose features meet the threshold for gay categorization compared with straight categorization, resulting in a small number of gay categorizations overall. However, we could not test this assumption directly in the previous studies because they involved morphed faces that lacked a true sexual orientation. Study 6 addressed this issue. We hypothesized that adaptation to hyper-masculine male faces would shift the range of features compelling gay categorizations, centering the perceptual threshold on the gender continuum and, therefore, reducing the straight categorization bias in judgments of real people.

### Method

**Participants.** There were 129 Internet users who completed the study ( $M_{age} = 36.38$ ; 60% women; 82% White; 91% straight).

**Stimuli.** Stimuli were identical to Study 5 with the addition of the 20 male faces described in Study 1.

**Procedure.** The procedure was similar to Study 5 with one addition: At pretest and posttest, participants judged the sexual orientation of the morphed faces from Study 5 as well as the real male faces from Study 1. Participants saw these two sets of faces in separate blocks, with block order counterbalanced across participants. Aside from the addition of real faces, all other procedures were identical to Study 5.

### Results and Discussion

Study 6 had two primary aims. First, we sought to replicate our previous findings about visual exposure as a mechanism by which perceivers calibrate the range of gendered facial features they associate with gay and straight categories. Second, we sought to test whether this recalibration alters the straight categorization bias documented above. Specifically, we tested whether visual adaptation to hyper-masculine faces reduced the magnitude of the straight categorization bias by altering the perceptual threshold for categorizing men as gay. We did so by fitting psychometric curves to the judgment data and calculating PSEs for the morphed faces and signal detection bias parameters for the real faces at pretest and posttest using the same analytic procedures described above.

We excluded 9 participants who failed a manipulation check testing their knowledge that half of the targets identified as gay, 49 participants who had nonconvergent or very poorly fitting curves at pretest, and 42 participants who had nonconvergent or very poorly fitting curves at posttest ( $R^2 < .60$ ).<sup>5</sup> As before, we intentionally recruited enough participants to ensure sufficient power after these exclusions. Indeed, after removing participants who failed the manipulation check and those with poorly fitting curves, we still had an effective sample of at least 70 participants. A power analysis revealed that 67 participants were required to achieve 80% power to detect small-to-medium effects ( $d = 0.35$ ; based on

Studies 4–5) in a dependent-samples  $t$  test when  $\alpha = .05$ . That said, the pattern of results replicates if we include participants who failed the manipulation check as well as in secondary analyses that included those with misfitting curves (see below).

#### Range of gendered cues supporting gay categorizations.

We first replicated our previous findings by examining the PSE for sexual orientation categorizations before and after adaptation to hyper-masculine faces. At pretest, the PSE was significantly below the midpoint of the gender continuum ( $M = -9.13$ ,  $SD = 33.63$ ),  $t(77) = -2.38$ ,  $p = .020$ ,  $d = 0.27$ , indicating there was a smaller range of gendered cues compelling gay categorizations than straight categorizations. At posttest, however, the PSE was not significantly different from the midpoint of the gender continuum ( $M = 2.69$ ,  $SD = 37.01$ ),  $t(84) = 0.67$ ,  $p = .505$ ,  $d = 0.07$ . The range of gendered cues compelling gay and straight categorizations was equivalent after adaptation to hyper-masculine faces.

To systematically test whether the PSE changed as a function of visual adaptation, we subtracted Pretest PSE from Posttest PSE. We then subjected this difference score to a one-sample  $t$  test. As expected, the PSE was significantly shifted toward the masculine end of the gender spectrum after adaptation to hyper-masculine faces ( $M = 10.29$ ,  $SD = 29.18$ ),  $t(69) = 2.95$ ,  $p = .004$ ,  $d = 0.35$ . These findings provide a direct replication of Study 5, indicating that the limited range of gendered cues supporting gay categorizations is modulated by exposure to highly gendered exemplars.

We conducted additional analyses to allay concerns about data exclusions in the PSE analysis. Specifically, we examined the probability of gay categorizations as a function of morph level, allowing us to include data from all participants who passed the manipulation check. Using GEEs to account for nested data, we regressed Sexual Orientation Categorization ( $0 = straight$ ,  $1 = gay$ ) onto Gender Morph ( $-100 = androgynous$  to  $100 = hyper-gendered$ ). At pretest, the slope for Gender Morph was negative and significant,  $B = -0.01$ ,  $SE < 0.01$ ,  $z = -10.29$ ,  $p < .001$ , indicating that participants applied a gender inversion heuristic whereby gay categorizations became more likely as faces became more androgynous. More important for our purposes is the intercept, which indicates the logit value for gay categorizations when Gender Morph = 0. The intercept value fell significantly below zero,  $B = -0.31$ ,  $SE = 0.06$ ,  $z = -4.98$ ,  $p < .001$ , indicating that participants categorized targets at the midpoint of the gender scale as straight even when the base rate constraint should have rendered these categorizations gay. At posttest, the slope for Gender Morph remained negative and significant,  $B = -0.02$ ,  $SE < 0.01$ ,  $z = -9.42$ ,  $p < .001$ , indicating that participants still applied a gender inversion heuristic. Critically, however, the intercept value was no longer significant,  $B = -0.02$ ,  $SE = 0.07$ ,  $z = -0.26$ ,  $p = .793$ . These findings corroborate those from the PSE analysis, revealing that visual adaptation to hyper-masculine faces resulted in a more balanced mapping of gendered features to sexual orientation categorizations.

**Response bias.** Next, we tested whether and how shifts in the perceptual threshold for gay categorizations tracked response bias for sexual orientation categorizations of real faces. We began by examining response bias at pretest. As expected, there was a

<sup>5</sup> There was substantial overlap in the participants with misfitting curves at pretest and posttest.

significant bias toward straight categorizations ( $M = 0.25$ ,  $SD = 0.32$ ),  $t(69) = 6.54$ ,  $p < .001$ ,  $d = 0.78$ . At posttest, however, there was no evidence of bias ( $M = -0.04$ ,  $SD = 0.49$ ),  $t(69) = -0.72$ ,  $p = .473$ ,  $d = 0.08$ .

To systematically test whether the bias parameter changed as a function of visual adaptation, we subtracted Pretest Bias from Posttest Bias. We then subjected this difference score to a one-sample  $t$  test. As expected, the straight categorization bias was significantly reduced after adaptation to hyper-masculine faces ( $M = -0.29$ ,  $SD = 0.38$ ),  $t(69) = -6.44$ ,  $p < .001$ ,  $d = 0.76$ .

In a final analysis, we used a multilevel mediation approach (Bauer, Preacher, & Gil, 2006) to test whether the change in response bias was accounted for by the change in PSE after adaptation. In this model, we specified Test Period ( $-0.5 = \text{pretest}$ ,  $0.5 = \text{posttest}$ ) as the predictor, Response Bias as the outcome, and PSE as the mediator. Results indicated a significant indirect effect of PSE, insofar as the confidence interval did not include zero: 95% confidence interval (CI) [0.01, 0.16]. Thus, the change in response bias observed as a function of visual adaptation was accounted for by the change in perceptual threshold for assigning gay category labels.

Study 6 extended our previous work in several ways. First, it replicated findings from Studies 4 and 5 indicating that there is a limited range of gendered cues supporting gay categorizations: Fewer faces had features consistent with perceivers' mental representation of the category gay relative to the category straight, despite the provision of information specifying equal numbers of gay and straight targets. Second, it replicated the finding from Study 5 that this range is calibrated on the basis of recent visual exposure: Adapting perceivers to hyper-masculine male faces altered the point of subjective equality for sexual orientation categorizations, balancing the range of features compelling straight and gay judgments. Third, and most importantly, it highlighted a causal link between this limited range of gendered cues and the straight categorization bias. Experimentally exposing perceivers to hyper-masculine male faces expanded the range of cues supporting gay categorizations, and in doing so, eradicated the straight categorization bias. Collectively, these findings highlight a perceptual mechanism underlying the straight categorization bias in sexual orientation judgments.

### General Discussion

Six studies provided the first systematic investigation of a response bias toward straight judgments in sexual orientation categorization. In Study 1, perceivers overutilized the straight category label even when they believed each trial had a 50% chance of revealing a gay target. Study 2 replicated this response bias among perceivers who received explicit information about the base rate of gay targets in a stimulus set, and Study 3 replicated it among perceivers who encountered stimulus sets with variable base rates of gay targets. Each of these initial studies provided some evidence that base rate matching contributes to the straight categorization bias. Indeed, perceivers tended to calibrate their judgments to the stated (Study 2) or presumed (Study 3) base rate for a stimulus set. That said, the straight categorizations bias persisted despite the provision of base rate information. Therefore, Studies 4–6 explored a second mechanism based upon perceivers' use of gendered cues when making sexual orientation judgments.

Study 4 revealed that the range of gendered cues compelling gay judgments is smaller than the range of gendered cues compelling straight judgments when perceivers believe the base rate of gay and straight targets is equal. Study 5 documented that perceptual exposure guides this imbalance: Using a visual adaptation paradigm, we found that exposing perceivers to hyper-gendered faces recentered their representational mapping of gendered features to sexual orientation categories, expanding the range of gendered features compelling gay categorizations. Finally, Study 6 linked this observation to our prior demonstrations of bias, showing that visual adaptation to hyper-gendered male faces reduces the magnitude of the straight categorization bias by centering threshold for gay categorizations.

This work provides new theoretical information about the processes and outcomes of sexual orientation categorization. First and foremost, our findings have implications for the ways researchers present findings related to sexual orientation categorization. The vast majority of scientific and popular discussion surrounding this topic has focused on the accuracy of perceivers' judgments. Although robust, the average magnitude of accuracy effects is modest ( $d = 0.29$ ; Tskhay & Rule, 2013) and often eclipsed by the average magnitude of response bias ( $d = 0.77$ ; Studies 1–3). Thus, focusing on accuracy without attending to response bias misrepresents the modal outcomes of sexual orientation categorization. Given that most perceivers achieve accurate sexual orientation categorizations by correctly judging straight targets to be straight, labeling the effect as “gaydar” is misleading. A more apt (albeit certainly less catchy) way of characterizing this literature would be “straight-dar.”

Our work also offers theoretical advances to researchers who study social categorization more broadly. One advance involves the role of base rates in social category judgments. Indeed, our initial studies showed that sexual orientation categorizations are at least somewhat sensitive to base rate information. Although we cannot deny the impact of base rates on judgments, it bears repeating that the straight categorization bias persisted despite the provision of base rate information. Indeed, the bias emerged even when restricting analyses to participants who passed a manipulation check testing their knowledge of the provided base rate. In this way, our findings mirror classic research showing that perceivers insufficiently account for base rates when forming judgments (Bar-Hillel, 1980; Kahneman & Tversky, 1973).

The current studies also provide additional nuance to discussions about the impact of base rates on social category judgments. As reported in Study 3, response bias was tethered more strongly to participants' presumed base rate than to the actual base rate of gay targets in a stimulus set. Furthermore, participants tended to infer a 50% base rate of gay targets regardless of their cumulative experience with the stimulus set. Unreported analyses from Studies 4–6 provided additional support for this conclusion. Indeed, participants' presumed base rates (as gleaned from their responses to the manipulation check) were significantly correlated with the number of gay categorizations they rendered for morphed faces in Study 4,  $r(271) = .40$ ,  $p < .001$ , for morphed faces at pretest in Study 5,  $r(174) = .17$ ,  $p = .074$ , and for real faces at pretest in Study 6,  $r(129) = .22$ ,  $p = .014$ . Thus, presumed base rates may play a stronger role than actual base rates in predicting categorizations for perceptually ambiguous social groups.

Of course, there is more work to be done to fully understand the impact of presumed base rates on the outcomes of sexual orientation categorization. Indeed, the few existing datasets that probe presumed base rates have yielded noticeably different results. For example, whereas most participants in Study 3 estimated a base rate of 50% gay targets, a recent Gallup poll found that typical Americans guess the base rate of lesbian/gay people in the population to be 25% (Morales, 2011). Perceivers' expectations about the commonality of lesbian and gay people appear to differ based upon sample characteristics as well as task demands (i.e., when asked about base rates in general vs. in a categorization study). Understanding whether and how perceivers integrate these personal beliefs into their social category judgments will be an important topic for future research. This is especially true because most studies of sexual orientation categorization overrepresent sexual minority targets relative to their frequency in the population, creating a mismatch between participants' beliefs about base rates in the population and their instructions for an empirical task. It might be the case that perceivers cannot overcome their prior beliefs, and instead calibrate judgments to a weighted average of the stated base rate and their beliefs about the population at large. We tested this possibility by restricting the sample from Study 3 to participants who estimated a base rate of 50%. If judgments reflect a weighted average of the presumed base rate for a categorization study and the Gallup estimate of a 25% presumed base rate for gay targets in the population, then perceivers should categorize 37.50% of the targets as gay. This was not the case: Participants categorized significantly fewer than 37.50% of the targets as gay ( $M = 28.86\%$ ,  $SD = 11.43\%$ ),  $t(144) = -9.10$ ,  $p < .001$ ,  $d = -0.76$ . This value does fall close to the Gallup estimate of 25%, though, suggesting that perceivers may encounter difficulty abandoning preexisting beliefs when they are provided with conflicting information. The current studies were not designed specifically to test this hypothesis, but we believe it represents an exciting avenue for future research examining how perceivers integrate preexisting knowledge with task instructions when forming social category judgments.

The above discussions highlight the fact that social categorizations are similar in form and function to many other decisions made under uncertainty. Like other decisions, perceivers are at least partially sensitive to base rate information when categorizing a concealable identity such as sexual orientation. At the same time, they do not adjust their judgments sufficiently, resulting in some notable decision biases. These findings suggest that judgment and decision-making may provide a useful framework for theorizing about social categorization. Continuing to probe the mechanisms of social categorization with methods from judgment and decision-making may help integrate these two areas of psychological inquiry.

Beyond their links to judgment and decision-making, the current studies also enrich our knowledge of the mechanisms guiding sexual orientation categorization. Whereas prior research consistently demonstrated that perceivers use a gender-inversion heuristic to categorize unknown others as gay (Freeman et al., 2010; Johnson et al., 2007; Lick et al., 2013; Stern et al., 2013), that work did not characterize the range of gendered cues compelling gay and straight judgments. Instead, the implicit assumption seemed to be that perceivers have a balanced range of cues supporting gay and straight categorizations,

because researchers have typically centered measures of gendered features and examined associations with sexual orientation judgments at 1 *SD* above and below the mean. The current studies lay bare this assumption and reveal that it does not accurately describe the categorization process, insofar as the range of cues supporting gay categorizations was notably smaller than the range of cues supporting straight categorizations. Our work, therefore, sheds new light on the gendered heuristics guiding sexual orientation categorization, which have been a topic of ongoing theoretical interest (see Cox et al., 2015). The field is now well poised to extend these insights, testing whether and how a limited range biases the outcomes of other social category judgments.

Finally, the current studies serve to join theories of visual adaptation, which have historically been examined within cognitive and vision science, with theories of social categorization, which have historically been examined within social science. Indeed, Study 5 provides the first known evidence that visual adaptation alters the perceptual threshold for gay categorization. In doing so, our work builds a foundation for new lines of research that examine how perceptual exposure impacts categorization along dimensions that were traditionally considered concealable. These insights join a new theoretical movement called *social vision*, which promotes the interdisciplinary merging of social and vision sciences to provide integrative knowledge about the perceptual processes underlying basic social phenomena (Johnson & Adams, 2013). Our visual adaptation findings also gesture toward potential real-world applications. Specifically, these findings suggest that increasing mainstream representation of masculine gay men and feminine lesbians may shift perceivers' mental representation of gendered cues in such a way that reduces their reliance on gendered features when categorizing sexual orientation. Because the gender-inversion heuristic represents an overgeneralization of features from the population to an individual, this process could be considered a method of reducing stereotypes related to sexual orientation. Of course, we caution that the current adaptation studies were restricted to tightly controlled laboratory situations; whether similar effects accrue after perceptual exposures in everyday life remains an important topic for future work.

We would be remiss if we did not note two limitations of the current work. First, as is true of almost any study of social perception, our stimuli were not fully representative of gay and straight men and women. Instead, the faces depicted relatively young White individuals displaying various emotional expressions. In Studies 1–3, most of the targets (70%) were smiling, though some displayed neutral expressions. In Studies 4–6, all of the targets displayed neutral expressions. These points are important given the burgeoning literature on intersectional social perception, which demonstrates that the dynamics of social categorization vary when multiple visual cues intersect (Cole, 2009; Johnson, Lick, & Carpinella, 2015). Although existing literature has consistently documented a straight categorization bias across different stimulus sets that included both men and women (Rule et al., 2008; Stern et al., 2013), diverse racial groups (Johnson & Ghavami, 2011; Rule et al., 2011), and various emotional expressions (Studies 1–3 vs. 4–6 of the current article), response bias may still vary as a function of unexamined target characteristics (e.g., eye gaze or emotional expression). Future researchers may



wish to consider how, when, and why the straight categorization bias differs across subpopulations of lesbians and gay men.

Second, our studies intentionally distorted the base rate of gay and lesbian targets relative to their true frequency or perceivers' beliefs about their frequency in the population. Recent demographic studies estimate a 3.5% base rate for sexual minorities in the population (Gates, 2011); in contrast, perceivers tend to estimate a base rate closer to 25% (Morales, 2011), and most of our studies had an empirical base rate of 50%. On a related note, Studies 4–6 presented participants with a series of similar-looking faces and required them to categorize targets given the constraint that half identified as gay. Thus, there may have been a mismatch between perceivers' beliefs about the frequency and appearance of sexual minorities and the targets they encountered. While it is common for scientists to constrain properties of the natural world to identify reliable patterns in social behavior, such constraints raise questions about the emergence of the straight categorization bias in more ecologically valid settings. In the future, it would be worthwhile to explore whether and how this bias emerges when empirical base rates approximate those observed in the general population.

In summary, the current studies highlight a pronounced tendency for observers to categorize unknown others as straight rather than gay. Investigating this response bias revealed new information about the perceptual mechanisms guiding sexual orientation categorization, and we believe it is likely to continue doing so in the future. As Helmholtz (1903) argued more than a century ago, "it is just those cases that are not in accordance with reality which are particularly instructive for discovering the laws of the processes by which normal perception originates." Decomposing judgment outcomes beyond basic accuracy can provide a fuller and more nuanced understanding of the decision-making processes underlying consequential forms of social categorization.

## References

- Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison Wesley.
- Ambady, N., Hallahan, M., & Conner, B. (1999). Accuracy of judgments of sexual orientation from thin slices of behavior. *Journal of Personality and Social Psychology, 77*, 538–547. <http://dx.doi.org/10.1037/0022-3514.77.3.538>
- Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica, 44*, 211–233. [http://dx.doi.org/10.1016/0001-6918\(80\)90046-3](http://dx.doi.org/10.1016/0001-6918(80)90046-3)
- Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods, 11*, 142–163. <http://dx.doi.org/10.1037/1082-989X.11.2.142>
- Bem, S. L. (1993). *The lenses of gender: Transforming the debate on sexual inequality*. New Haven, CT: Yale University Press.
- Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. In R. S. Wyer, Jr., (Ed.), *Stereotype activation and inhibition: Advances in social cognition* (Vol. 11, pp. 1–52). Mahwah, NJ: Erlbaum.
- Bodenhausen, G. V., & Peery, D. (2009). Social categorization and stereotyping in vivo: The VUCA challenge. *Social and Personality Psychology Compass, 3*, 133–151. <http://dx.doi.org/10.1111/j.1751-9004.2009.00167.x>
- Cole, E. R. (2009). Intersectionality and research in psychology. *American Psychologist, 64*, 170–180. <http://dx.doi.org/10.1037/a0014564>
- Cox, W. T. L., Devine, P. G., Bischmann, A. A., & Hyde, J. S. (2015). Inferences about sexual orientation: The roles of stereotypes, faces, and the gaydar myth. *Journal of Sex Research, 53*, 157–171.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5–18. <http://dx.doi.org/10.1037/0022-3514.56.1.5>
- Dyer, R. (1999). Stereotyping. In L. Gross & J. D. Woods (Eds.), *The Columbia reader: On lesbians and gay men in media, society, and politics* (pp. 297–301). New York, NY: Columbia University Press.
- Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review, 112*, 912–931. <http://dx.doi.org/10.1037/0033-295X.112.4.912>
- Estes, W. K. (1976). The cognitive side of probability learning. *Psychological Review, 83*, 37–64. <http://dx.doi.org/10.1037/0033-295X.83.1.37>
- Federal Bureau of Investigation. (2011). *Hate crime statistics, 2011*. Washington, DC: U.S. Department of Justice.
- 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. <http://dx.doi.org/10.1177/0146167210378755>
- Gates, G. J. (2011). *How many people are lesbian, gay, bisexual, and transgender?* Retrieved from <http://williamsinstitute.law.ucla.edu/wp-content/uploads/Gates-How-Many-People-LGBT-Apr-2011.pdf>
- Gaudio, R. P. (1994). Sounding gay: Pitch properties in the speech of gay and straight men. *American Speech, 69*, 30–57. <http://dx.doi.org/10.2307/455948>
- Hart, K. R. (2000). Representing gay men on American television. *The Journal of Men's Studies, 9*, 59–79. <http://dx.doi.org/10.3149/jms.0901.59>
- Hatzenbuehler, M. L. (2009). How does sexual minority stigma "get under the skin"? A psychological mediation framework. *Psychological Bulletin, 135*, 707–730. <http://dx.doi.org/10.1037/a0016441>
- Helmholtz, H. (1903). *Popular lectures on scientific subjects* (E. Atkinson, Trans.). New York, NY: Green.
- Johnson, K. L., & Adams, R. B., Jr. (2013). Social vision: An introduction. *Social Cognition, 31*, 633–635. <http://dx.doi.org/10.1521/soco.2013.31.6.633a>
- Johnson, K. L., & Ghavami, N. (2011). At the crossroads of conspicuous and concealable: What race categories communicate about sexual orientation. *PLoS ONE, 6*, e18025. <http://dx.doi.org/10.1371/journal.pone.0018025>
- Johnson, K. L., Gill, S., Reichman, V., & Tassinari, L. G. (2007). Swagger, sway, and sexuality: Judging sexual orientation from body motion and morphology. *Journal of Personality and Social Psychology, 93*, 321–334. <http://dx.doi.org/10.1037/0022-3514.93.3.321>
- Johnson, K. L., Lick, D. J., & Carpinella, C. M. (2015). Emergent research in social vision: An integrated approach to the determinants and consequences of social categorization. *Social and Personality Psychology Compass, 9*, 15–30. <http://dx.doi.org/10.1111/spc3.12147>
- Johnson, K. L., & Tassinari, L. G. (2005). Perceiving sex directly and indirectly: Meaning in motion and morphology. *Psychological Science, 16*, 890–897. <http://dx.doi.org/10.1111/j.1467-9280.2005.01633.x>
- Johnson, K. L., & Tassinari, L. G. (2007). Compatibility of basic social perceptions determines perceived attractiveness. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 5246–5251. <http://dx.doi.org/10.1073/pnas.0608181104>
- Johnston, L., Miles, L., & Macrae, N. (2010). Male or female? An investigation of factors that modulate the visual perception of another's sex. In E. Balceris & G. Lassiter (Eds.), *The social psychology of visual perception* (pp. 103–118). East Sussex: Psychology Press.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*, 237–251. <http://dx.doi.org/10.1037/h0034747>



- Katz-Wise, S. L., & Hyde, J. S. (2012). Victimization experiences of lesbian, gay, and bisexual individuals: A meta-analysis. *Journal of Sex Research, 49*, 142–167. <http://dx.doi.org/10.1080/00224499.2011.637247>
- Kite, M. E., & Deaux, K. (1987). Gender belief systems: Homosexuality and the implicit inversion theory. *Psychology of Women Quarterly, 11*, 83–96. <http://dx.doi.org/10.1111/j.1471-6402.1987.tb00776.x>
- Lick, D. J., Durso, L. E., & Johnson, K. L. (2013). Minority stress and physical health among sexual minorities. *Perspectives on Psychological Science, 8*, 521–548. <http://dx.doi.org/10.1177/1745691613497965>
- Lick, D. J., & Johnson, K. L. (2013). Fluency of visual processing explains prejudiced evaluations following categorization of concealable identities. *Journal of Experimental Social Psychology, 49*, 419–425. <http://dx.doi.org/10.1016/j.jesp.2013.01.001>
- Lick, D. J., & Johnson, K. L. (2014a). Perceptual underpinnings of antigay prejudice: Negative evaluations of sexual minority women arise on the basis of gendered facial features. *Personality and Social Psychology Bulletin, 40*, 1178–1192. <http://dx.doi.org/10.1177/0146167214538288>
- Lick, D. J., & Johnson, K. L. (2014b). Recalibrating gender perception: Face aftereffects and the perceptual underpinnings of gender-related biases. *Journal of Experimental Psychology: General, 143*, 1259–1276. <http://dx.doi.org/10.1037/a0034516>
- Lick, D. J., Johnson, K. L., & Gill, S. V. (2013). Deliberate changes to gendered body motion influence basic social perceptions. *Social Cognition, 31*, 656–671. <http://dx.doi.org/10.1521/soco.2013.31.6.656>
- Lick, D. J., Johnson, K. L., & Gill, S. V. (2014). Why do they have to flaunt it? Perceptions of communicative intent explain antigay prejudice based upon brief exposure to nonverbal cues. *Social Psychological and Personality Science, 5*, 927–935. <http://dx.doi.org/10.1177/1948550614537311>
- Linville, S. E. (1998). Acoustic correlates of perceived versus actual sexual orientation in men's speech. *Folia Phoniatrica et Logopaedica, 50*, 35–48. <http://dx.doi.org/10.1159/000021447>
- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*, 1122–1135. <http://dx.doi.org/10.3758/s13428-014-0532-5>
- Martin, D., & Macrae, C. N. (2007). A face with a cue: Exploring the inevitability of person categorization. *European Journal of Social Psychology, 37*, 806–816. <http://dx.doi.org/10.1002/ejsp.445>
- Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence. *Psychological Bulletin, 129*, 674–697.
- Morales, L. (2011). U.S. adults estimate that 25% of Americans are gay or lesbian. Retrieved from <http://www.gallup.com/poll/147824/adults-estimate-americans-gay-lesbian.aspx>
- Munson, B., & Babel, M. (2007). Loose lips and silver tongues, or, projecting sexual orientation through speech. *Language and Linguistics Compass, 1*, 416–449. <http://dx.doi.org/10.1111/j.1749-818X.2007.00028.x>
- Olivola, C. Y., & Todorov, A. (2010). Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology, 46*, 315–324. <http://dx.doi.org/10.1016/j.jesp.2009.12.002>
- Raley, A. B., & Lucas, J. L. (2006). Stereotype or success? Prime-time television's portrayals of gay male, lesbian, and bisexual characters. *Journal of Homosexuality, 51*, 19–38. [http://dx.doi.org/10.1300/J082v51n02\\_02](http://dx.doi.org/10.1300/J082v51n02_02)
- Rendall, D., Vasey, P. L., & McKenzie, J. (2008). The Queen's English: An alternative, biosocial hypothesis for the distinctive features of "gay speech". *Archives of Sexual Behavior, 37*, 188–204. <http://dx.doi.org/10.1007/s10508-007-9269-x>
- 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. <http://dx.doi.org/10.1007/s10508-008-9405-2>
- Rule, N. O., & Ambady, N. (2008). Brief exposures: Male sexual orientation is accurately perceived at 50-ms. *Journal of Experimental Social Psychology, 44*, 1100–1105. <http://dx.doi.org/10.1016/j.jesp.2007.12.001>
- Rule, N. O., Ambady, N., Adams, R. B., Jr., & Macrae, C. N. (2008). Accuracy and awareness in the perception and categorization of male sexual orientation. *Journal of Personality and Social Psychology, 95*, 1019–1028. <http://dx.doi.org/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. <http://dx.doi.org/10.1016/j.jesp.2009.07.010>
- Rule, N. O., Ishii, K., Ambady, N., Rosen, K. S., & Hallett, K. C. (2011). Found in translation: Cross-cultural consensus in the accurate categorization of male sexual orientation. *Personality and Social Psychology Bulletin, 37*, 1449–1507. <http://dx.doi.org/10.1177/0146167211415630>
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments & Computers, 31*, 137–149. <http://dx.doi.org/10.3758/BF03207704>
- Stern, C., West, T. V., Jost, J. T., & Rule, N. O. (2013). The politics of gaydar: Ideological differences in the use of gendered cues in categorizing sexual orientation. *Journal of Personality and Social Psychology, 104*, 520–541. <http://dx.doi.org/10.1037/a0031187>
- Tabak, J. A., & Zayas, V. (2012). The roles of featural and configural face processing in snap judgments of sexual orientation. *PLoS ONE, 7*, e36671. <http://dx.doi.org/10.1371/journal.pone.0036671>
- Tskhay, K. O., & Rule, N. O. (2013). Accuracy in categorizing perceptually ambiguous groups: A review and meta-analysis. *Personality and Social Psychology Review, 17*, 72–86. <http://dx.doi.org/10.1177/1088868312461308>
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 453–458. <http://dx.doi.org/10.1126/science.7455683>
- Webster, M. A. (2012). Evolving concepts of sensory adaptation. *F1000 Biology Reports, 4*, 21. <http://dx.doi.org/10.3410/B4-21>
- Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428*, 557–561. <http://dx.doi.org/10.1038/nature02420>
- Wilder, D. A. (1986). Social categorization: Implications for creation and reduction of intergroup bias. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 293–355). New York, NY: Academic Press.
- Zeger, S. L., & Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics, 42*, 121–130. <http://dx.doi.org/10.2307/2531248>
- Zhao, C., Serriès, P., Hancock, P. J. B., & Bednar, J. A. (2011). Similar neural adaptation mechanisms underlying face gender and tilt aftereffects. *Vision Research, 51*, 2021–2030. <http://dx.doi.org/10.1016/j.visres.2011.07.014>

Received February 12, 2015

Revision received March 10, 2016

Accepted March 14, 2016 ■