# Evidence for the Absence of Stimulus Quality Differences in Tests of the Accuracy of Sexual Orientation Judgments: A Reply to Cox et al. (2016)

Nicholas O. Rule

University of Toronto

rule@psych.utoronto.ca

Kerri L. Johnson

University of California-Los Angeles

kerri.johnson@ucla.edu

Jonathan B. Freeman

New York University

jon.freeman@nyu.edu

Abstract

A paper recently published in this journal (Cox, Devine, Bischmann, & Hyde, 2016) questioned the validity of existing research on the accurate judgment of sexual orientation from photographs of faces. Specifically, those authors reported a confound in their stimuli whereby the photos of sexual minorities (gay men and lesbians) were of higher quality than the photos of heterosexuals. Based on this finding, they concluded that the accuracy in judging sexual orientation from faces demonstrated in the broader literature is an artifact of stimulus quality differences. Here, we addressed this claim by systematically testing the numerous photo sets that we have used in 61 studies documenting accurate judgments of sexual orientation from facial cues published since 2007. Contrary to their claim, the overwhelming majority of studies (93%) showed no significant differences in photo quality according to sexual orientation. Of those that did show differences, most revealed that heterosexual targets' photos were actually of *higher* quality than sexual minority targets' photos. In addition, we highlight additional research using stimuli equated for quality differences overlooked in the recent article, lending further support to the conclusion that sexual orientation is legible from multiple sensory cues.

# Evidence for the Absence of Stimulus Quality Differences in Tests of the Accuracy of Sexual Orientation Judgments: A Reply to Cox et al. (2016)

Recently, Cox, Devine, Bischmann, and Hyde (2016) questioned the well-documented observation that people can infer others' sexual orientation from observing cues expressed through nonverbal behaviors. Popularly dubbed "gaydar," dozens of studies published since 1987 have demonstrated that people can judge men's and women's sexual orientation significantly better than what would be expected from chance based upon their speech, movement, and visible aspects of their appearance. Cox et al. have disputed these findings and suggested that research on the ability to judge sexual orientation from indirect cues perpetuates stereotypes about sexual minorities. This possibility, if correct, deserves careful consideration. Unfortunately, aspects of Cox et al.'s evidence reveal substantial errors in the logic leading to this conclusion. The goal of this commentary is therefore to clarify what past studies have shown about people's performance in judging sexual orientation from nonverbal cues and to provide critical tests of the central claims made by Cox et al. regarding the methodological integrity of the relevant past work.

First, Cox et al. (2016) focused specifically on the role that facial appearance plays in people's judgments of sexual orientation. Beginning in 2007, roughly 22 published articles have reported data showing that individuals could guess others' sexual orientation significantly better than chance when viewing photos of their faces or individual facial features (see Tskhay & Rule, 2013, for review). In addition, numerous other studies have examined the accuracy of judgments of sexual orientation from cues other than the face. For example, Ambady, Hallahan, and Conner (1999) asked research participants to estimate the Kinsey score of men and women based on very brief video clips (as short as 1 second). They found that the participants' judgments significantly correlated with the targets' own self-reported scores. These studies did not rely directly on information in the targets' faces and, in some conditions, degraded facial information so that it was barely visible. Johnson and colleagues likewise examined judgments of men's and women's sexual orientation from figural outlines and point-light displays of their body movements (Johnson, Gill, Reichman, & Tassinary, 2007; Lick, Johnson, & Gill, 2013). Critically, their studies showed no facial cues at all, yet found levels of accuracy similar to those reported in Ambady et al.'s work. Finally, absent visual perception altogether, several studies have shown that voice and speech patterns alone can reveal information about one's sexual orientation (e.g., Sulpizio et al., 2015), and also that judgments from these cues are often more accurate than judgments from visual cues (see Tskhay & Rule, 2013, for formal comparison). Moreover, there is a long history of research documenting gender nonconformity in the appearance and behaviors of sexual minorities (see LeVay, 2011; Rule, in press; and Tskhay & Rule, 2015a, for reviews), and recent work linking morphological differences in appearance to accurate perceptions of sexual orientation (e.g., Skorska, Geniole, Vrysen, McCormick, & Bogaert, 2015; Valentova, Kleisner, Havlíček , & Neustupa, 2014). Thus, sexual orientation appears to be legible from a variety of different nonverbal channels beyond those specific to the face. These findings present a considerable challenge to Cox et al.'s conclusion that gaydar is a "myth." However, these studies were absent from Cox et al.'s review of the literature.

There is, of course, also a large body of evidence that people can perceive sexual orientation from facial cues. The authors of these studies examined targets of different sexes (men and women; e.g., Tabak & Zayas, 2012), races (Asian, Black, and Caucasian; Johnson & Ghavami, 2011; Rule, 2011), and national cultures (the Czech Republic, Japan, Spain, and the U.S.; Rule, Ishii, Ambady, Rosen, & Hallett, 2011; Valentova & Havlíček, 2013). They also explored how individual perceivers' levels of accuracy may vary according to their own race, sex, and culture; their own sexual orientation (Johnson & Ghavami, 2011; Rule, Ambady, Adams, & Macrae, 2007), their fertility status (Rule, Rosen, Slepian & Ambady, 2011), their familiarity with sexual minorities (Brambilla, Riva, & Rule, 2013), their political ideology (Stern, West, Jost, & Rule, 2013), and their level of anti-gay prejudice (Rule et al., 2015). In spite of this convergent evidence, Cox et al. (2016) suggested that these findings may be invalid because the set of stimuli they created for their own research exhibited a confound between sexual orientation and photo quality. Specifically, according to their described methods, Cox et al. used photos of heterosexual and sexual minority individuals obtained from Internet personal advertisements, an approach that was common in early published works. Although they found that participants in their study could categorize these targets as gay, straight, and lesbian significantly better than chance guessing, they also found that the photos posted online by the sexual minority targets were perceived as significantly higher in quality than those posted by the heterosexuals. When they selected subsets of photos from each group with equal quality scores (notably far fewer than most other studies testing the legibility of sexual orientation from faces), differences in judgments of the targets' sexual orientation were no longer significant. Cox et al. then claimed that this confound between sexual orientation and photo quality plagues all prior stimulus sets used in the literature. However, this generalization naturally commits a syllogistic leap, as it assumes that the photos used by Cox et al. are representative of those used by other researchers. Here, we show that they are not.

Cox et al. (2016) used photos from online personal advertisements, a reasonable approach given historical norms. Many of the past studies estimating the accuracy of sexual orientation judgments from faces used photos downloaded from websites and, specifically, from personal advertisements. However, many have not. For instance, Tabak and Zayas (2012) did not use photos from personal advertisements, but from Facebook.com. Similarly, Rule and Ambady (2008) and Rule, Ambady, Adams, and Macrae (2008) also used photos from both personal advertisements and from Facebook.com. Stern et al. (2013) and Stern, West, Jost, and Rule (2014) used photos of gay and heterosexual men taken under standardized conditions in a laboratory (though they did not report the details of this in their papers), as did Valentova and colleagues (Valentova & Havlíček, 2013; Valentova et al., 2014). Cox et al.'s generalization is not relevant to these studies.

Beyond this, there is also reason for skepticism that Cox et al.'s (2016) confound in stimulus quality is characteristic of stimulus sets in other studies using photos from personal advertisements. All photos used by Rule et al. (2008), for example, were intentionally degraded. Color information was removed, the size of the photos was made uniform, and—most important—the photos were cropped to show only individual facial features (see also Rule, Ambady, & Hallett, 2009, and Tskhay, Feriozzo, & Rule, 2013). Moreover, the *conditions* of judgment were additionally degraded in other studies, such that participants saw the photos as briefly as 40 milliseconds yet continued to demonstrate levels of accuracy statistically equivalent to those under longer viewing times (Rule & Ambady, 2008; Rule et al., 2009). Although these experimental methods suggest that the accuracy of sexual orientation judgments from facial cues is quite robust, it remains possible that photo quality might vary in these studies, thereby contributing to the observed accuracy in judgments. To test this claim, we conducted a series of studies in which we assessed the photo quality of the images that the current authors have collectively used in our published studies testing the accuracy of sexual orientation judgments from photos of faces since 2007.

## Method

Cox et al. (2016) implied that the extant literature has used only a single stimulus set in testing the legibility of sexual orientation from the face. This characterization is incorrect. Indeed, between 2007 and 2015, our published research testing the accuracy of sexual orientation judgments utilized 13 distinct sets of photos sourced from the Internet (not including several additional sets from other sources). Twelve of these stimulus sets consisted of photos drawn from dating websites; the thirteenth consisted of photos drawn from Facebook.com (see Table 1). This was not acknowledged in Cox et al.'s description of these studies.

Following the procedures described by Cox et al. (2016), we asked participants to rate the photo quality of the images of the sexual minority (gay, lesbian, and bisexual) and heterosexual targets from 1 (*Very Poor*) to 7 (*Excellent*). We aimed to recruit 30 participants to rate the photos in each condition (i.e., photo set) in order to match the sample size used by Cox et al. when they evaluated the quality of their male photos. Thus, we requested 390 Mechanical Turk Workers to rate one of the 13 sets of photos, though a total of 468 participants actually completed the study (78 participants did not claim payment and thus did not count towards the sample we requested from Amazon). We excluded an additional 63 participants either because their Internet Protocol address appeared twice (suggesting that they might have repeated the study; thus excluding the second iteration but retaining the first; $n$ = 21), because they did not complete the full study ($n$ = 33), or because they failed to answer an attention check question correctly ($n$ = 9); see Table 1 for sample demographics by condition.

Given that our research question concerned the quality of the targets' photos, we treated the participants as trials and the targets as the unit of analysis, thereby aggregating the participants' judgments to form average quality ratings for each target. Prior to doing so, we calculated the inter-rater agreement for the participants in each sample to confirm the suitability of averaging their scores together (all Cronbach's αs ≥ .93). We then compared the quality ratings for the sexual minority (gay, lesbian, and bisexual) versus heterosexual targets in each photo set, Bonferroni correcting for multiple comparisons. Results of a power analysis based on the effect size that Cox et al. (2016) obtained in their tests of photo quality ($M_{\text{Cohen's } d} = 0.80$) indicated that we would need groups of 42 sexual minority and heterosexual targets, respectively, to achieve 5% false-positive and false-negative rates in a two-tailed independent samples t-test. Twelve of our stimulus sets met this threshold; the one remaining stimulus set that did not (Set 4 in Tables 1 and 2) included only 40 images per group and resulted in 94% power.

In addition to measuring the photo quality of the targets in each of our 13 Internet-sourced photo sets, we also wanted to confirm that sub-selections of the photos taken from these 13 sets that we used in our published studies (including cropped and transformed versions) did not differ in photo quality. Specifically, we have measured the accuracy of judgments of sexual orientation from faces in a total of 61 studies spanning 17 published articles since 2007; 50 of these studies used images of faces from online sources (47 from dating websites and three from Facebook.com, spanning 15 published articles) whereas the remaining 11 studies used photos of gay and heterosexual male targets' faces photographed under identical standardized conditions in the laboratory (spanning three published articles). To determine whether our stimuli in these studies were confounded as Cox et al. (2016) suggested, we tested the individual subsets used in each of these 61 studies for differences in photo quality. Because we could not commit the resources needed to recruit human participants to separately rate all of the 6,121 images contributing to these 61 studies, we used a publicly-available image quality assessment algorithm developed by computer scientists to quantify photo quality (Mittal, Moorthy, & Bovik, 2012). As the algorithm simulates the quality assessment of human judges, we validated its efficacy by correlating the scores with the photo quality judgments provided by our participants for the 13 superordinate photo sets prior to comparing the quality ratings for sexual minority and heterosexual targets.

## Results

As listed in Table 2, eight of the 13 photo sets showed no significant difference in quality between the sexual minority and heterosexual targets (regardless of whether we corrected for multiple comparisons; Bonferroni-corrected $\alpha = .004$). Four out of five of the remaining photo sets showed differences *opposite* those reported by Cox et al. (2016); that is, the straight targets' photos were higher in quality than the sexual minority targets' photos. Non-significant effects can sometimes cumulatively reveal important patterns in data, however. To address this possibility, we meta-analytically combined the effect sizes from each of these analyses with those calculated from the quality comparisons reported by Cox et al. ($r = .51$ for male faces and $r = .21$ for female faces) by converting the $r$-values to Fisher's $z$ scores so that we could aggregate them and subject them to inferential tests. Doing so produced a mean effect size opposite the direction found by Cox et al. that did not significantly differ from 0, $M_z = -.06$, $SD = .29$, 95% CI [-.20, .09].[1]

[1] Alternatively, using a multilevel model that accounted for the 13 different photo sets and modeled the time at which the participants made their ratings as a covariate (because we did not collect the data for all photo sets in a single study) showed that photo quality did not differ between the sexual minority and heterosexual targets, $B = .022$, $SE = .029$, 95% CI [-.034, .078], $t(1271.21) = 0.77$, $p = .44$ (degrees of freedom calculated using Satterthwaite approximation).

Because we used photos from these sets across numerous studies testing the accuracy of sexual orientation judgments, we thought it prudent to investigate the extent to which quality differences might have emerged in the studies using the exact stimuli tested in each condition of each study we published. We accomplished this using a computer algorithm that objectively estimates humans' subjective quality assessments in which lower scores indicate higher quality (i.e., less visual noise). Consistent with the demonstrated performance of the algorithm in its development (Mittal et al., 2012), the subjective opinions of our participants correlated well with the quality scores produced by the algorithm ($M_z = -.71$, $SD = .28$, 95% CI [-.87, -.56]; see Table 2 for correlations according to each photo set). We therefore tested the image quality of the photos used in each of the 61 studies that we have collectively published from 2007 until the time of present writing (see Supplementary Table 1 online for summary). Results showed that nine of the 61 photo sets significantly differed at traditionally accepted levels ($\alpha < .05$) but only four of these significantly differed when correcting for multiple comparisons ($\alpha < .0008$). Of these four, only three were in the same direction as those found by Cox et al. (2016), and they were all from a redundant stimulus set (comparisons of the gay and bisexual male vs. heterosexual faces used in Studies 1A, 1B, and 2 of Ding & Rule's, 2012, report). Aggregating the effect sizes for the comparisons of the heterosexual versus sexual minority targets across all of these studies, but removing redundant stimulus sets to counter inflation (final $k = 42$), showed a small mean effect size close to 0 that was not significant, $M_z = -.02$, $SD = .20$, 95% CI [-.08, .05].

## Discussion

Cox et al. (2016) reported in this journal that disparities in the quality of the photos they used in their studies explained the accuracy of judgments of sexual orientation, not only in their own research but also in the broader literature. Specifically, they found that images of gay men and lesbians they obtained from Internet dating website advertisements were of higher perceived quality than images they collected of heterosexual people. This observation of their own stimuli led Cox et al. to suggest that all studies testing the accuracy of sexual orientation judgments for images culled from the Internet exhibit the same flaw. Here, we have demonstrated that this assumption is unfounded.

Across 13 sets of stimuli, used in parcel in 61 studies spanning a total of 17 published articles for 4,154 unique images in total, we found no systematic evidence that photos of sexual minorities are of higher quality than photos of heterosexuals. Using Cox et al.'s (2016) own method of measuring photo quality, we found that sexual minority targets had higher quality photos than straight targets in only one of the 13 stimulus sets used in prior research. Moreover, eight of the 13 stimulus sets showed no significant differences in photo quality between heterosexual and sexual minority targets, and the remaining four stimulus sets significantly differed in the *opposite* direction: photo quality was higher for heterosexual than sexual minority targets. Aggregating across all 13 photo sets, the mean effect size of the significant and non-significant differences showed a tendency for heterosexual targets to have higher quality photos than sexual minority targets, though this average effect was not reliable. Thus, the photos that we have collectively used, which constitute the preponderance of research examining the accuracy of judging sexual orientation from facial photographs, do not share the stimulus confound present in Cox et al.'s investigation.

These results from participants' subjective ratings were corroborated by an "objective" computer-based method of measuring photo quality. After quantifying photo quality using a computer algorithm, we found that 93% (i.e., 57) of the 61 studies that we have published since

2007 showed no differences in photo quality as a function of the targets' sexual orientation. Of the remaining four studies, one showed differences in the opposite direction of those reported by Cox et al. (2016); the three that did show consistent differences were drawn from a single set of stimuli used in three separate studies published within one paper. When combining the effects of all of the comparisons of the 42 non-redundant stimulus sets, we again observed a small overall effect in the opposite direction that did not approach statistical significance.

Considering these findings, Cox et al.'s (2016) conclusion that sexual orientation cannot be accurately judged from photos of the face is unsubstantiated. The confound observed in Cox et al.'s stimuli does not characterize the majority of stimulus sets used in research demonstrating that sexual orientation can be judged from static facial cues. Equally important, Cox et al. failed to account for the number of studies using stimuli from non-dating website sources, such as photos taken under standardized conditions in the laboratory, Internet-based photos carefully sampled from sources other than personal advertisements, and judgments from video and audio recordings that typically show stronger effects (see Tskhay & Rule, 2013). Cox et al.'s general implication that "gaydar" does not exist therefore is not supported by the broad survey of empirical data on the subject.

This should not be surprising, in light of numerous studies that have demonstrated *genuine* differences between sexual minority and heterosexual individuals in their appearance and behavior (e.g., Rieger, Linsenmeier, Gygax, & Bailey, 2008; Skorska et al., 2015; Valentova et al., 2014). Based on these real differences, a volume of studies examining perceptions of sexual orientation across a variety of channels of expression and communication have shown evidence for greater than chance accuracy. These studies go far beyond face perception to show that sexual orientation can be accurately discerned in many contexts ranging from various cues in body movement (e.g., Johnson et al., 2007), expressive style (e.g., Ambady et al., 1999; Rieger, Linsenmeier, Gygax, Garcia, & Bailey, 2010), and the voice (e.g., Rendall, Vasey, & McKenzie, 2008). These modalities not only yield higher levels of accuracy in judgments of sexual orientation than cues extracted from the face (Tskhay & Rule, 2013), but are also immune to the potential confound proposed by Cox et al. (2016).

This is not to say that this area of research is without limitations. For instance, most of the research on perceptions of sexual orientation has been conducted by a small group of researchers and there is evidence that this has introduced some bias into the literature. In their meta-analysis, Tskhay and Rule (2013) noted that much of the research related to judgments of perceptually ambiguous groups (including sexual orientation) came from studies done by Rule and his colleagues. Their tests showed that studies published by Rule and his collaborators showed smaller effects than those published by other research groups, thus lowering the mean overall effect size describing the magnitude of sexual orientation's legibility.

Another obvious question concerns the representativeness of the photos of heterosexuals and sexual minorities used in these studies. Indeed, as some of the research reviewed here has already shown, appearance cues related to sexual orientation vary across social contexts and geographic regions, as do the heuristics that perceivers use to judge them. Yet most of the stimuli used to test the accuracy of judging sexual orientation from faces have relied on a single medium: photos downloaded from online dating websites. Although online dating is increasingly common, many people do not date online and those that do might have particular motivations or intentions that influence which photos they select to post for this purpose (e.g., Ellison, Heino, & Gibbs, 2006; Finkel, Eastwick, Karney, Reis, & Sprecher, 2012). Thus, such photos may or may not represent the broader population. Other studies used photos from different sources of online social media (i.e., Facebook.com). Whereas these photos were likely not directly influenced by the targets' self-presentation biases (i.e., they consisted of photos posted by the targets' friends and with individuals other than the targets also in the photos; see Rule et al., 2008), other factors might have limited their representativeness; for example, not everyone uses Facebook.com. Finally, additional studies photographed individuals recruited to the lab (e.g., Stern et al., 2013). Here, self-presentational biases would not have limited representativeness, particularly as the targets did not know they would be photographed until arriving. Although we have experienced that nearly all individuals agree to participate as targets, such samples are limited to people participating in psychology studies (usually university students enrolled in psychology courses). Thus, each approach has drawbacks. However, the convergent results drawn from using these various sources of stimuli (including those examined across diverse racial and cultural groups) suggests a shared underlying property that allows for the reliable perception of sexual orientation. Nevertheless, future research would benefit from more comprehensively exploring the contextual and environmental factors that may influence both the perception and expression of appearance-based cues to sexual orientation.

Indeed, observers must necessarily exploit some aspect of visual stimuli when making judgments of sexual orientation. If not photo quality, as Cox et al. (2016) claimed, what visual information might support these judgments? Several studies have found that gender atypicality is one important cue used by observers: gay men tend to exhibit more feminine facial appearances than heterosexual men, and lesbian women tend to exhibit more masculine facial appearances than heterosexual women (Freeman, Johnson, Ambady, & Rule, 2010; Lyons, Lynch, Brewer, & Bruno, 2014; Skorska et al., 2015; Valentova et al., 2014). Other research showed that heterosexual men have more symmetrical faces than gay men (Hughes & Bremme, 2011).

Tskhay and Rule (2015b) additionally found that subtle differences between gay and heterosexual men's emotional expressions explain sexual orientation judgments beyond the contributions of gender atypicality and facial symmetry. Thus, stereotypes about differences in heterosexual and gay targets' facial masculinity and femininity might not only guide perceivers' judgments (as noted by Cox et al.) but, on average, may also reflect a kernel of truth.

References

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

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

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

Ding, J. Y. C., & Rule, N. O. (2012). Gay, straight, or somewhere in between: Accuracy and bias in the perception of bisexual faces. *Journal of Nonverbal Behavior, 36*, 165–176.

Ellison, N., Heino, R., & Gibbs, J. (2006). Managing impressions online: Self-presentation processes in the online dating environment. *Journal of Computer-Mediated Communication, 11*, 415-441.

Finkel, E. J., Eastwick, P. W., Karney, B. R., Reis, H. T., & Sprecher, S. (2012). Online dating: A critical analysis from the perspective of psychological science. *Psychological Science in the Public Interest, 13,* 3-66.

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.

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

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

Johnson, K. L., Gill, S., Reichman, V., & Tassinary, L. G. (2007). Swagger, sway, and sexuality: Judging sexual orientation from body motion and morphology. *Journal of Personality and Social Psychology,* 93*,* 321–334.

LeVay, S. (2011). *Gay, Straight, and the Reason Why: The Science of Sexual Orientation.* New York, NY: Oxford University Press.

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.

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

Mittal, A., Moorthy, A. K., & Bovik, A. C. (2012). No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing, 12,* 4695-4708.

Remedios, J. D., Chasteen, A. L., Rule, N. O., & Plaks, J. E. (2011). Impressions at the intersection of ambiguous and obvious social categories: Does gay + Black = likable? *Journal of Experimental Social Psychology, 47*, 1312–1315.

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.

Rieger, G., Linsenmeier, J. A., Gygax, L., & Bailey, J. M. (2008). Sexual orientation and childhood gender nonconformity: evidence from home videos. *Developmental Psychology*, *44*, 46-58.

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.

Rule, N. O. (2011). The influence of target and perceiver race in the categorization of male sexual orientation. *Perception, 40*, 830–839.

Rule, N. O. (in press). Perceptions of sexual orientation from minimal cues. *Archives of Sexual Behavior.*

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.

Rule, N. O., Ambady, N., Adams, R. B., Jr., & Macrae, C. N. (2007). Us and them: Memory advantages in perceptually ambiguous groups. *Psychonomic Bulletin & Review, 14,* 687-692.

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.

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.

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.

Rule, N. O., Macrae, C. N., & Ambady, N. (2009). Ambiguous group membership is extracted automatically from faces. *Psychological Science, 20*, 441–443.

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

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

Rule, N. O., Tskhay, K. O., Freeman, J. B., & Ambady, N. (2014). On the interactive influence of facial appearance and explicit knowledge in social categorization. *European Journal of Social Psychology, 44*, 529-535.

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

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.

Stern, C., West, T. V., Jost, J. T., & Rule, N. O. (2014). "Ditto heads": Do conservatives perceive greater consensus within their ranks than liberals? *Personality and Social Psychology Bulletin, 40*, 1162-1177.

Sulpizio, S., Fasoli, F., Maass, A., Paladino, M. P., Vespignani, F., Eyssel, F., & Bentler, D. (2015). The sound of voice: Voice-based categorization of speakers' sexual orientation within and across languages. *PLoS ONE, 10,* e0128882.

Tabak, J. A., & Zayas, V. (2012). The roles of featural and configural face processing in snap judgments of sexual orientation. *PloS ONE, 7*, e36671.

Tskhay, K. O., Feriozzo, M. M., & Rule, N. O. (2013). Facial features influence the categorization of female sexual orientation. *Perception, 42,* 1090-1094.

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.

Tskhay, K. O., & Rule, N. O. (2015a). Sexual orientation across culture and time. In S. Safdar & N. Kosakowska-Berzecka (Eds.), *The Psychology of Gender and Culture* (pp. 55-73). New York, NY: Springer.

Tskhay, K. O., & Rule, N. O. (2015b). Emotions facilitate the communication of ambiguous group memberships. *Emotion*, *15,* 812-826.

Valentova, J. V., & Havlíček, J. (2013). Perceived sexual orientation based on vocal and facial stimuli is linked to self-rated sexual orientation in Czech men. *PLoS ONE, 8,* e82417.

Valentova, J. V., Kleisner, K., Havlíček, J., & Neustupa, J. (2014). Shape differences between the faces of homosexual and heterosexual men. *Archives of Sexual Behavior, 43,* 353-361.

.

Table 1

*Target and Rater Descriptive Statistics According to Condition within Original Source Article for Each Photo Set*

| | | | $n$ Targets | | Raters | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Sex | | | Age | | |
| Photo Set | Original Source Article | Condition | SM | H | *N* | α | $n_{Male}$ | $n_{Female}$ | $n_{Other}$ | *M* | *SD* | $n_H$ |
| 1 | Rule & Ambady (2008)[a] | Dating Websites | 45 | 45 | 30 | .94 | 21 | 9 | 0 | 36.57 | 11.88 | 28 |
| 2 | Rule & Ambady (2008) | Facebook.com | 80 | 80 | 44 | .98 | 19 | 25 | 0 | 37.70 | 13.73 | 42 |
| 3 | Rule, Ambady, & Hallett (2009) | Studies 1-3 | 94 | 98 | 32 | .95 | 13 | 19 | 0 | 37.59 | 14.14 | 28 |
| 4 | Freeman et al. (2010) | Study 3 | 40 | 40 | 39 | .97 | 14 | 24 | 1 | 34.15 | 11.86 | 36 |
| 5 | Johnson & Ghavami (2011) | - | 150 | 150 | 31 | .98 | 11 | 20 | 0 | 36.65 | 11.68 | 28 |
| 6 | Rule (2011) | Asian | 50 | 50 | 32 | .93 | 13 | 19 | 0 | 39.59 | 12.11 | 30 |
| 7 | Rule (2011) | Black | 50 | 50 | 34 | .97 | 13 | 20 | 1 | 33.09 | 10.44 | 32 |
| 8 | Rule (2011) | Caucasian | 50 | 50 | 35 | .97 | 18 | 17 | 0 | 37.74 | 12.60 | 33 |
| 9 | Rule, Ishii, et al. (2011) | Japanese | 50 | 50 | 41 | .98 | 19 | 22 | 0 | 34.46 | 10.97 | 37 |
| 10 | Rule, Ishii, et al. (2011) | Spanish | 51 | 49 | 37 | .98 | 17 | 20 | 0 | 36.05 | 11.25 | 31 |
| 11 | Ding & Rule (2012) | Male | 86 | 44 | 39 | .97 | 13 | 26 | 0 | 35.26 | 12.12 | 36 |
| 12 | Ding & Rule (2012) | Female | 80 | 40 | 44 | .97 | 20 | 24 | 0 | 34.73 | 11.17 | 41 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 13 | Rule, Ambady, & Hallett (2009) | MC | 99 | 98 | 30 | .98 | 13 | 17 | 0 | 35.90 | 9.48 | 29 |

*Note.* SM = sexual minority, H = heterosexual, α = Cronbach's alpha inter-rater reliability coefficient, *M* = mean, *SD* = standard deviation, MC = Manipulation Check. [a] Rule et al. (2007) used a subset of these faces in an earlier publication.

Table 2

*Descriptive Statistics, Significance Tests, and Correlations with the Objective Quality Ratings for the Mean Subjective Quality Ratings Provided by the Participants for Each of the Distinct Parent Photo Sets*

| | SM | | H | | | | | | $M_{\text{Difference}}$ 95% CI | | Correlation with Objective Measure | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | 95% CI | |
| Photo Set | *M* | *SD* | *M* | *SD* | *t* | df | *p* | $r_{\text{ES}}$ | LL | UL | *r* | LL | UL |
| 1 | 4.14 | 0.87 | 4.15 | 1.00 | -0.08 | 88 | .94 | -.01 | -0.41 | 0.38 | -.69 | -.80 | -.55 |
| 2 | 3.96 | 1.16 | 3.87 | 0.99 | 0.54 | 158 | .59 | .04 | -0.24 | 0.43 | -.74 | -.80 | -.66 |
| 3 | 3.74 | 0.89 | 4.20 | 0.67 | -4.04 | 173.19[a] | < .001 | -.29 | -0.68 | -0.24 | -.37 | -.49 | -.24 |
| 4 | 4.07 | 0.88 | 4.21 | 0.82 | -0.72 | 78 | .48 | -.08 | -0.51 | 0.24 | -.41 | -.56 | -.23 |
| 5 | 4.30 | 1.09 | 3.76 | 1.15 | 4.16 | 298 | < .001 | .23 | 0.28 | 0.79 | -.62 | -.69 | -.55 |
| 6 | 3.66 | 0.89 | 3.51 | 0.69 | 0.91 | 98 | .36 | .09 | -0.17 | 0.46 | -.36 | -.55 | -.14 |
| 7 | 3.99 | 1.10 | 3.79 | 1.11 | 0.92 | 98 | .36 | .09 | -0.24 | 0.64 | -.59 | -.70 | -.48 |
| 8 | 3.75 | 1.02 | 4.11 | 0.90 | -1.86 | 98 | .067 | -.18 | -0.74 | 0.025 | -.63 | -.75 | -.48 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9 | 3.28 | 1.00 | 4.03 | 0.99 | -3.76 | 98 | < .001 | -.36 | -1.15 | -0.35 | -.70 | -.79 | -.60 |
| 10 | 3.87 | 1.21 | 3.93 | 1.30 | -0.23 | 98 | .82 | -.02 | -0.56 | 0.44 | -.82 | -.88 | -.74 |
| 11 | 4.19 | 1.03 | 4.44 | 0.74 | -1.44 | 128[a] | .15 | -.13 | -0.60 | 0.10 | -.36 | -.50 | -.21 |
| 12 | 3.76 | 1.02 | 4.31 | 0.69 | -3.46 | 107.20[a] | .001 | -.32 | -0.86 | -0.23 | -.51 | -.63 | -.37 |
| 13 | 3.07 | 1.10 | 4.45 | 0.92 | -9.54 | 195 | < .001 | -.56 | -1.67 | -1.10 | -.83 | -.87 | -.79 |

*Note.* SM = sexual minority, H = heterosexual, *M* = mean, *SD* = standard deviation, df = degrees of freedom, ES = effect size, CI = confidence interval, LL = lower limit, UL = upper limit; [a] Degrees of freedom corrected for heteroscedasticity.

**Supplementary Table 1**

Descriptive statistics and mean comparisons for computer-based quality scores of images of sexual minority versus heterosexual targets in sexual orientation judgment accuracy studies published by the authors since 2007

| | | | $n_{\text{Targets}}$ | | Quality Ratings | | | | Comparisons of Means | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | SM | | H | | | | | | 95% CI of Mean Differences | |
| Article | Study | Condition | SM | H | *M* | *SD* | *M* | *SD* | *t* | *df* | *p* | $r_{ES}$ | LL | UL |
| Photos from Dating Websites | | | | | | | | | | | | | | |
| Rule et al. (2007) | - | - | 40 | 40 | 34.56 | 8.24 | 36.91 | 10.37 | -1.12 | 78 | .27 | -.13 | -6.52 | 1.82 |
| Rule & Ambady (2008)[1] | 1a | - | 45 | 45 | 34.08 | 7.84 | 37.80 | 10.35 | -1.92 | 88 | .06 | -.20 | -7.57 | 0.12 |
| Rule et al. (2008)[2] | 1 | - | 45 | 36 | 34.25 | 7.93 | 36.75 | 10.36 | -1.23 | 79 | .22 | -.14 | -6.55 | 1.54 |
| Rule et al. (2008)[2] | 2 | Control | 45 | 36 | 34.25 | 7.93 | 36.75 | 10.36 | -1.23 | 79 | .22 | -.14 | -6.55 | 1.54 |
| Rule et al. (2008) | 2 | Eyes Occluded | 45 | 36 | 36.96 | 8.18 | 41.35 | 9.77 | -2.20 | 79 | .03 | -.24 | -8.36 | -0.43 |
| Rule et al. (2008) | 2 | Hair Removed | 45 | 36 | 56.54 | 7.06 | 59.57 | 11.78 | -1.36 | 54.46[a] | .18 | -.18 | -7.50 | 1.43 |
| Rule et al. (2008) | 2 | Mouth Occluded | 45 | 36 | 34.17 | 7.68 | 37.71 | 9.92 | -1.81 | 79 | .07 | -.20 | -7.43 | 0.36 |
| Rule et al. (2008) | 3 | Salient Features Removed | 45 | 45 | 66.08 | 4.82 | 68.07 | 9.34 | -1.27 | 65.85[a] | .21 | -.15 | -5.12 | 1.14 |
| Rule et al. (2008) | 3 | Eyes Only | 45 | 45 | 64.74 | 13.67 | 61.39 | 11.62 | 1.25 | 88 | .21 | .13 | -1.97 | 8.67 |
| Rule et al. (2008)[3] | 3 | Hair Only | 45 | 45 | 59.13 | 10.23 | 65.39 | 11.41 | -2.74 | 88 | .01 | -.28 | -10.80 | -1.72 |
| Rule et al. (2008)[4] | 3 | Mouth Only | 45 | 45 | 70.10 | 5.53 | 71.45 | 8.02 | -0.93 | 78.15[a] | .36 | -.10 | -4.24 | 1.54 |
| Rule et al. (2008) | 4 | Eyes Only | 45 | 45 | 64.74 | 13.67 | 61.39 | 11.62 | 1.25 | 88 | .21 | .13 | -1.97 | 8.67 |
| Rule et al. (2008)[3] | 4 | Hair Only | 45 | 45 | 59.13 | 10.23 | 65.39 | 11.41 | -2.74 | 88 | .01 | -.28 | -10.80 | -1.72 |
| Rule et al. (2008)[4] | 4 | Mouth Only | 45 | 45 | 70.10 | 5.53 | 71.45 | 8.02 | -0.93 | 78.15[a] | .36 | -.10 | -4.24 | 1.54 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Rule et al. (2008)[1] | 4 | Full Face | 45 | 45 | 34.08 | 7.84 | 37.80 | 10.35 | -1.92 | 88 | .06 | -.20 | -7.57 | 0.12 |
| Rule, Ambady, & Hallett (2009)[5] | 1a | - | 94 | 98 | 22.45 | 8.32 | 20.54 | 8.26 | 1.59 | 190 | .11 | .11 | -0.45 | 4.27 |
| Rule, Ambady, & Hallett (2009) | 1b | - | 94 | 98 | 56.32 | 20.22 | 52.33 | 20.59 | 1.34 | 190 | .18 | .10 | -1.87 | 9.85 |
| Rule, Ambady, & Hallett (2009)[6] | 2 | - | 94 | 98 | 29.29 | 6.62 | 29.91 | 5.94 | -0.69 | 190 | .49 | -.05 | -2.41 | 1.17 |
| Rule, Ambady, & Hallett (2009)[6] | 3 | - | 94 | 98 | 29.29 | 6.62 | 29.91 | 5.94 | -0.69 | 190 | .49 | -.05 | -2.41 | 1.17 |
| Rule, Ambady, & Hallett (2009) | MC | - | 99 | 98 | 59.87 | 19.16 | 23.97 | 15.28 | 14.55 | 186.57[a] | .00 | .73 | 31.03 | 40.77 |
| Rule, Macrae, & Ambady (2009) | - | - | 10 | 10 | 35.37 | 7.72 | 36.93 | 8.99 | -0.42 | 18 | .68 | -.10 | -9.44 | 6.31 |
| Freeman et al. (2010) | 2 | - | 60 | 60 | 45.00 | 15.96 | 47.17 | 13.97 | -0.79 | 118 | .43 | -.07 | -7.56 | 3.25 |
| Freeman et al. (2010) | 3 | - | 40 | 40 | 29.17 | 10.01 | 24.92 | 11.03 | 1.81 | 78 | .08 | .20 | -0.44 | 8.94 |
| Johnson & Ghavami (2011) | - | - | 150 | 150 | 38.93 | 9.68 | 41.20 | 10.57 | -1.94 | 298 | .05 | -.11 | -4.57 | 0.04 |
| Remedios et al. (2011) | 1 | - | 52 | 52 | 34.45 | 12.95 | 32.92 | 12.75 | 0.61 | 102 | .54 | .06 | -3.46 | 6.53 |
| Rule (2011) | 1 | - | 150 | 150 | 35.75 | 11.58 | 35.30 | 11.94 | 0.34 | 298 | .74 | .02 | -2.21 | 3.13 |
| Rule (2011) | 2 | Asian | 50 | 50 | 36.57 | 12.15 | 39.21 | 8.46 | -1.26 | 87.47[a] | .21 | -.13 | -6.80 | 1.53 |
| Rule (2011) | 2 | Black | 50 | 50 | 37.93 | 9.98 | 37.70 | 12.71 | 0.10 | 98 | .92 | .01 | -4.31 | 4.76 |
| Rule (2011) | 2 | Caucasian | 50 | 50 | 32.76 | 12.06 | 28.98 | 11.74 | 1.59 | 98 | .12 | .16 | -0.94 | 8.51 |
| Rule, Ishii, et al. (2011)[1] | - | American | 45 | 45 | 34.08 | 7.84 | 37.80 | 10.35 | -1.92 | 88 | .06 | -.20 | -7.57 | 0.12 |
| Rule, Ishii, et al. (2011) | - | Japanese | 50 | 50 | 33.75 | 13.96 | 33.89 | 15.90 | -0.04 | 98 | .97 | .00 | -6.07 | 5.81 |
| Rule, Ishii, et al. (2011) | - | Spanish | 51 | 49 | 28.69 | 13.80 | 29.56 | 16.10 | -0.29 | 98 | .77 | -.03 | -6.82 | 5.07 |
| Rule, Rosen, et al. (2011)[7] | 1 | - | 40 | 40 | 31.81 | 11.45 | 30.79 | 11.26 | 0.40 | 78 | .69 | .05 | -4.04 | 6.07 |
| Rule, Rosen, et al. (2011)[5] | 2 | - | 94 | 98 | 22.45 | 8.32 | 20.54 | 8.26 | 1.59 | 190 | .11 | .11 | -0.45 | 4.27 |
| Rule, Rosen, et al. (2011)[7] | 3a | - | 40 | 40 | 31.81 | 11.45 | 30.79 | 11.26 | 0.40 | 78 | .69 | .05 | -4.04 | 6.07 |
| Rule, Rosen, et al. (2011)[5] | 3b | - | 94 | 98 | 22.45 | 8.32 | 20.54 | 8.26 | 1.59 | 190 | .11 | .11 | -0.45 | 4.27 |
| Ding & Rule (2012)[8] | 1 | - | 86 | 44 | 29.96 | 9.86 | 37.63 | 10.40 | -4.12 | 128 | <.001 | -.34 | -11.35 | -3.99 |
| Ding & Rule (2012)[8] | 2a | - | 86 | 44 | 29.96 | 9.86 | 37.63 | 10.40 | -4.12 | 128 | <.001 | -.34 | -11.35 | -3.99 |
| Ding & Rule (2012)[8] | 2b | - | 86 | 44 | 29.96 | 9.86 | 37.63 | 10.40 | -4.12 | 128 | <.001 | -.34 | -11.35 | -3.99 |
| Ding & Rule (2012) | 3 | - | 80 | 40 | 23.30 | 9.70 | 17.56 | 7.71 | 3.27 | 118 | <.001 | .29 | 2.26 | 9.23 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Brambilla et al. (2013)[1] | - | - | 45 | 45 | 34.08 | 7.84 | 37.80 | 10.35 | -1.92 | 88 | .06 | -.20 | -7.57 | 0.12 |
| Tskhay et al. (2013) | - | Full Face | 48 | 48 | 22.14 | 9.47 | 19.26 | 8.76 | 1.54 | 94 | .13 | .16 | -0.82 | 6.57 |
| Tskhay et al. (2013) | - | Internal Features | 48 | 48 | 36.36 | 5.52 | 35.87 | 6.25 | 0.41 | 94 | .68 | .04 | -1.90 | 2.88 |
| Tskhay et al. (2013) | - | Eyes Only | 48 | 48 | 39.84 | 9.13 | 40.18 | 8.07 | -0.19 | 94[a] | .85 | -.02 | -3.83 | 3.16 |
| Rule et al. (2014)[9] | 1 | - | 10 | 10 | 32.77 | 8.47 | 37.54 | 10.51 | -1.12 | 18 | .28 | -.26 | -13.73 | 4.21 |
| Rule et al. (2014)[9] | 2 | - | 10 | 10 | 32.77 | 8.47 | 37.54 | 10.51 | -1.12 | 18 | .28 | -.26 | -13.73 | 4.21 |
| Tskhay & Rule (2015)[1] | 4a | - | 45 | 45 | 34.08 | 7.84 | 37.80 | 10.35 | -1.92 | 88 | .06 | -.20 | -7.57 | 0.12 |
| Photos from Facebook.com | | | | | | | | | | | | | | |
| Rule et al. (2008) | 5 | All Stimuli | 80 | 80 | 34.55 | 11.05 | 35.37 | 10.95 | -0.47 | 158 | .64 | -.04 | -4.25 | 2.62 |
| Rule et al. (2008) | 5 | Selected by Others | 69 | 64 | 35.01 | 11.00 | 35.41 | 11.14 | -0.21 | 131 | .83 | -.02 | -4.21 | 3.39 |
| Rule & Ambady (2008) | 1b | - | 80 | 80 | 46.46 | 8.84 | 52.41 | 9.53 | -2.04 | 158 | .04 | -.16 | -5.83 | -0.09 |
| Photos Taken in the Lab | | | | | | | | | | | | | | |
| Stern et al. (2013)[10] | 1 | - | 15 | 15 | 4.12 | 7.11 | 4.57 | 7.29 | -0.17 | 28 | .87 | -.03 | -5.84 | 4.93 |
| Stern et al. (2013)[10] | 2 | - | 15 | 15 | 4.12 | 7.11 | 4.57 | 7.29 | -0.17 | 28 | .87 | -.03 | -5.84 | 4.93 |
| Stern et al. (2014)[10] | 1 | - | 15 | 15 | 4.12 | 7.11 | 4.57 | 7.29 | -0.17 | 28 | .87 | -.03 | -5.84 | 4.93 |
| Stern et al. (2014)[10] | 2 | - | 15 | 15 | 4.12 | 7.11 | 4.57 | 7.29 | -0.17 | 28 | .87 | -.03 | -5.84 | 4.93 |
| Stern et al. (2014)[10] | 3 | - | 15 | 15 | 4.12 | 7.11 | 4.57 | 7.29 | -0.17 | 28 | .87 | -.03 | -5.84 | 4.93 |
| Tskhay & Rule (2015) | 2a | Gay-Acting | 24 | 24 | 10.81 | 7.95 | 9.13 | 4.39 | 0.90 | 46 | .37 | .13 | -2.06 | 5.41 |
| Tskhay & Rule (2015) | 2a | Straight-Acting | 24 | 24 | 8.85 | 7.00 | 9.01 | 4.23 | -0.10 | 46 | .92 | -.01 | -3.53 | 3.20 |
| Tskhay & Rule (2015) | 2a | Neutral | 24 | 24 | 8.88 | 5.83 | 9.38 | 3.41 | -0.36 | 46 | .72 | -.05 | -3.28 | 2.27 |
| Tskhay & Rule (2015) | 3a | Angry Morph | 20 | 20 | 12.74 | 6.23 | 12.51 | 3.59 | 0.15 | 38 | .89 | .02 | -3.02 | 3.49 |
| Tskhay & Rule (2015) | 3a | Happy Morph | 20 | 20 | 11.77 | 5.91 | 11.72 | 3.44 | 0.04 | 38 | .97 | .01 | -3.04 | 3.15 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Tskhay & Rule (2015) | 3a | Neutral | 20 | 20 | 9.54 | 5.96 | 9.19 | 3.37 | 0.23 | 38 | .82 | .04 | -2.75 | 3.44 |

***Note:*** Matching numerical superscripts indicate identical stimulus sets. [a] Degrees of freedom corrected for heteroscedasticity.

**Abbreviations:**

SM = sexual minority (i.e., gay, lesbian, or bisexual)
H= heterosexual
*M* = mean
*SD* = standard deviation
df = degrees of freedom
ES = effect size
CI = confidence interval
LL = lower limit
UL = upper limit
MC = manipulation check