Routledge
Taylor & Francis Group

Journal of Homosexuality

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/wjhm20

Who Has a Better Auditory Gaydar? Sexual Orientation Categorization by Heterosexual and Lesbian, Gay, and Bisexual People

Fabio Fasoli, Anne Maass & Luna Berghella

To cite this article: Fabio Fasoli, Anne Maass & Luna Berghella (2023) Who Has a Better Auditory Gaydar? Sexual Orientation Categorization by Heterosexual and Lesbian, Gay, and Bisexual People, Journal of Homosexuality, 70:5, 876-899, DOI: 10.1080/00918369.2021.2004796

To link to this article: https://doi.org/10.1080/00918369.2021.2004796

JOURNAL OF HOMOSEXUALITY
2023, VOL. 70, NO. 5, 876–899
https://doi.org/10.1080/00918369.2021.2004796

Routledge
Taylor & Francis Group

Who Has a Better Auditory Gaydar? Sexual Orientation Categorization by Heterosexual and Lesbian, Gay, and Bisexual People

Fabio Fasoli, PhD [a,b], Anne Maass, PhD [c], and Luna Berghella, MSc [c]

[a]School of Psychology, University of Surrey, Guildford, UK; [b]Centro de Investigação E Intervenção Social, Instituto Universitário de Lisboa (Iscte-iul), Lisbon, Portugal; [c]Dipartimento Di Psicologia Dello Sviluppo E Della Socializzazione, Università Di Padova, Padua, Italy

ABSTRACT

Lesbian, gay and bisexual (LGB) people are supposed to be better at gaydar than heterosexual. Across two studies we examined auditory gaydar performed by LGB and heterosexual listeners. In Study 1 participants (n = 127) listened to male and female speakers (n = 10) and judged their sexual orientation on a binary choice (gay/lesbian vs. heterosexual). In Study 2, participants (n = 192) judged speakers' (n = 31) sexual orientation on a Kinsey-like scale (1 = exclusively heterosexual, 7 = exclusively gay/lesbian). Results showed gaydar judgments differences in relative terms that did not indicate an overall gaydar accuracy. Moreover, LGB participants were not better at gaydar than heterosexual participants but rather showed a shift in criterion when making auditory gaydar judgments, namely they report a weaker straight categorization bias. Overall, these findings contribute to the understanding of sexual orientation categorization among heterosexual majority and LGB minority groups.

KEYWORDS

Gaydar; voice; sexual orientation; gender typicality

Is s/he gay or straight? You should be able to tell! There is a general assumption among people that lesbian, gay and bisexual (LGB) people are better at detecting other gay men and lesbian women as they have a more developed "gaydar." Originally, the term "gaydar" defined "an inexplicable intuition seemingly innately bestowed upon members of the gay community" (Woolery, 2007, p. 11). Shelp (2003) suggests that gaydar has an adaptative function that allows LGB individuals to recognize each other and, thus, form a community. Literature has shown that both heterosexual and LGB individuals categorize others' sexual orientation (SO) but has not yet come to a conclusion on whether LGB individuals outperform heterosexual individuals in gaydar accuracy. The present research examines this question with a focus on auditory gaydar, namely SO judgments based on vocal cues.

CONTACT Fabio Fasoli f.fasoli@surrey.ac.Uk School of Psychology, University of Surrey, Stag Hill, University Campus, Guildford GU2 7XH, UK.

Supplemental data for this article can be accessed on the publisher's website.

Communicating sexual orientation

Historically, LGB people have used different ways to communicate their SO to other ingroup members. In the USA, asking a man whether he was a "friend of Dorothy" or wearing a red tie were ways to signal a gay SO to other gay men (Chauncey, 1994). In the 60s-70s, some British gay men communicated through "Polari," a code language allowing them to communicate without being understood by heterosexual people (Baker, 2003). Nowadays, some gay men and lesbian women may signal their SO in different ways. For instance, they may use explicit cues such as clothing or adornments to express their SO (Clarke, Hayfield, & Huxley, 2012), and modify consciously or unconsciously their speech after coming out or when talking with gay friends (Daniele, Fasoli, Antonio, Sulpizio, & Maass, 2020; Podesva, 2007). Not only that, gay men and lesbian women also signal their own and detect others' SO by using direct and prolonged eye-gaze (Nicholas, 2004). Such behaviors often go hand in hand with specific gestures, postures and facial expressions that guide gaydar judgments (Barton, 2015). Hence, LGB people are particularly engaged in signaling and recognizing SO, possibly as a way to avoid isolation and share experiences of stigmatization (Shelp, 2003). In line with common beliefs, this may allow them to develop superior skills to accurately capture others' SO.

Who has a better auditory gaydar?

Gaydar judgments can be difficult and may vary depending on the cue under consideration (Fasoli & Maass, 2018). Rule (2017) has suggested that an above chance gaydar accuracy (varying between 60–65%) can be found in the literature. Such accuracy may be lower when voice is the only available cue. Indeed, results on auditory gaydar accuracy are mixed, with some studies supporting accuracy (Gaudio, 1994; Linville, 1998; Tracy, Bainter, & Satariano, 2015; Valentova & Havlíček, 2013) and others suggesting a biased and largely inaccurate distinction of gay- from heterosexual-sounding speakers that occurs regardless of the way the speakers self-identified (Munson, 2007; Smyth, Jacobs, & Rogers, 2003; Sulpizio et al., 2020, 2015). Individuals are generally hesitant to label others as gay (Sulpizio et al., 2015) and gaydar judgments seem to be biased by a default assumption that everyone is heterosexual unless gender-atypical cues are present, the so-called *straight categorization bias* (Lick & Johnson, 2016). Indeed, perceived gender typicality is linked to SO judgments (Kachel, Simpson et al., 2018; Kachel, Steffens, Preuß, & Simpson, 2020). When voice is concerned, gay men's voice quality influences the attribution of masculinity (see Ravenhill & de Visser, 2017) and perception of speakers as sounding gay (lesbian) and feminine (masculine) go hand in hand (Munson, 2007). Speakers themselves have expectations on whether they will be perceived as gay/lesbian or heterosexual depending on

how gender typical they sound (Fasoli, Hegarty, Maass, & Antonio, 2018). The role of gender stereotyping in gaydar research has been shown by Cox, Devine, Bischmann, and Hyde (2016), who found that gaydar judgments are influenced by stereotypes and that believing gaydar is accurate led individuals to use stereotypes as SO diagnostic cues. In fact, gay "stereotypical" voices increase the perceived differences between speakers' SO (Kachel et al., 2018).

Literature comparing gaydar across LGB and heterosexual people is scarce independent of what kind of cue (visual, auditory) one considers. Seminal work on this topic by Berger, Hank, Rauzi, and Simkins (1987), who asked LGB and heterosexual individuals to judge targets' SO on the basis of short video interviews, demonstrated that accuracy rates were consistently below chance levels for both LGB and heterosexual participants. Subsequent studies found a tendency for gay men to be better at gaydar than heterosexual men (Shelp, 2003) and for lesbian women to be more accurate when judging female targets (Ruben, Hill, & Hall, 2014). Research considering both target gender and SO (Ambady et al.,1999) showed an overall accuracy (above chance) of gaydar judgments that was higher for female targets when visual stimuli were involved. Moreover, LGB participants outperformed heterosexual participants. Other research on visual gaydar confirmed a better accuracy for female targets (Brewer & Lyons, 2016; Lyons et al., 2014), but did not find LGB participants to be more accurate than heterosexual participants. Instead, LGB individuals were generally more likely to label targets as gay/lesbian. This could be explained by the fact that they found this labeling as less stigmatizing than heterosexual individuals did (see Alt, Lick, & Johnson, 2020). Thus, in signal detection terms, the difference between LGB and heterosexual participants may not so much lie in their capacity to discriminate gay from heterosexual targets, but in their response bias. In line with signal detection theory, we refer here to this bias as "criterion" guiding gaydar judgments.

None of the above studies comparing LGB and heterosexual gaydar has investigated voice. Rieger, Linsenmeier, Gygax, Garcia, and Bailey (2010) examined voice while considering multiple cues (e.g., movement, appearance) either together or separately. Results showed an overall gaydar accuracy that occurred particularly when participants watched the full video or focused on targets' speech and appearance. However, results were not consistent in showing that LGB were better than heterosexual participants. Indeed, lesbian women were less accurate than others in judging men's SO based on movement and speech, but LGB participants were better than heterosexual participants in judging female targets' SO from movements. Valentova and Havlíček (2013), found instead that GB men were better than heterosexual women at correctly judging SO of male targets from their faces and voices. Hence, mixed

results exist and until now, no other research has investigated gaydar differences between LGB and heterosexual people when voice is the only cue available.

Voice is an interesting channel through which SO may be communicated. Voice is often overlooked as a source of information although, unbeknownst to most people, it tends to exert a particularly strong effect on social categorization (Rakić, Steffens, & Mummendey, 2011). Also, and most importantly, voices are extremely variable (Kreiman, Keating, Park, Rastifar, & Alwan, 2015). They not only change over time (Lavan, Burton, Scott, & McGettigan, 2019), but also depending on the contexts (for SO, see Daniele et al., 2020). Given the great variability of voice across contexts, familiarity becomes a critical variable in recognition accuracy. As far as the recognition of individual voices is concerned, Lavan et al. (2019, p. 95) have argued that "overall, familiarity with a speaker or a stimulus provides listeners with an advantage in identity processing. In contrast to unfamiliar listeners, familiar listeners appear to be able to better generalize the information from familiar signals (e.g., speech in a familiar language) to less familiar signals (e.g., laughter or speech in another language)." With regards to SO, Brambilla, Riva, and Rule (2013) have shown that familiarity with gay men is associated with more accurate face-based gaydar judgments. It is possible that LGB individuals have more experience with their ingroup members and are more familiar with SO cues, including voice, than heterosexual people. If the advantage of familiarity in identity processing (*does this voice belong to person A?)* also holds for categorical processing (*does this voice belong to a gay speaker?)*, then one could predict that LGB participants are more accurate in recognizing others' SO from voice alone. This prediction is based on the assumption that, compared to heterosexual participants, LGB members have considerably more experience with gay/lesbian voices and are exposed to larger and more varied samples of gay/lesbian voices, that they experience across multiple contexts. Thus, they may be better at detecting intra-group variation occurring among gay/lesbian individuals (see Kachel et al., 2018). In contrast, exposure to heterosexual voices should be similar for LGB and heterosexual individuals given that the majority of the population identifies as heterosexual. If this argument is correct, then LGB individuals may indeed be more accurate in identifying others' SO, as popular wisdom suggests. But what does accuracy really mean?

Different facets of accuracy

Accuracy is a muddy concept that can take on different meanings in gaydar research. One definition refers to the capacity to *differentiate* between the two SO categories of speakers. In this case, it is sufficient that listeners judge heterosexual speakers as more heterosexual than gay/lesbian speakers. Thus,

accuracy is defined in *relative* terms and reflects the ability to distinguish the two categories of speakers. However, this does not mean that speakers are correctly classified. For instance, as found in many studies (e.g., Munson, 2007; Sulpizio et al., 2020, 2015; Valentova & Havlíček, 2013), both heterosexual and gay/lesbian speakers may be identified as "heterosexual," but to different degrees.

A different and more stringent definition refers to accuracy in an *absolute* sense, such that a single speaker is placed in the correct category (in case of binary decisions) or on the correct side of the spectrum (in case of interval Kinsey-like scales). In the case of multiple speakers, it requires that the majority of speakers is placed in the correct category or on the correct side of the spectrum. It is therefore possible that people are accurate in relative terms, such that they are able to reliably distinguish the two categories of speakers, but that they are inaccurate in absolute terms because they perceive all speakers as heterosexual (see *straight categorization bias*, Lick & Johnson, 2016).

To complicate things further, absolute accuracy may either reflect a superior ability to differentiate gay/lesbian from heterosexual speakers or a shift in criterion (or both). We define criterion here as the more or less restrictive tendency to label others as gay/lesbian (also named "response bias" in previous research). Take the example of ratings reported by heterosexual and LGB participants on a 7-point Kinsey-like scale with one extreme representing "exclusively heterosexual" and the other "exclusively gay/lesbian" (first row of Figure 1). Heterosexual participants may differentiate between the two SO categories as shown by statistically different means (second row of Figure 1). Still, both means are on the "heterosexual pole" of the scale indicating that heterosexual participants assume all speakers to be heterosexual, although to

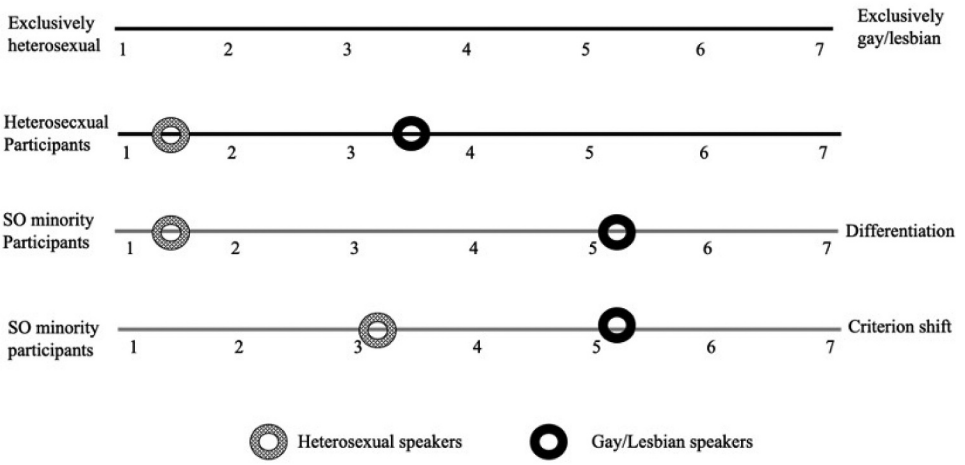

Figure 1. Increased accuracy driven by better differentiation vs. criterion shift.

different degrees, and suggesting a strong straight categorization bias (Lick & Johnson, 2016). LGB participants may be more accurate in classifying gay/lesbian speakers, as common wisdom suggests. However, their gaydar judgments could take different forms. LGB participants could outperform heterosexual participants because of a better ability to differentiate between the two categories, hence they perceived gay/lesbian speakers as clearly gay while maintaining a perception of heterosexual speakers as heterosexual (both means on the correct side of the scale, third row of Figure 1). Another possibility is that the superior performance of LGB participants is simply a side-effect of a different criterion (forth row of Figure 1). In this case, the perceived distance between the two categories is exactly the same as that indicated by heterosexual participants (second row, Figure 1), but the values are shifted toward the right side of the continuum, resulting in greater accuracy. An analogous argument can be made for dichotomous responses (*gay or straight?)* using signal detection theory that allows to consider both correct differentiation (accuracy) and shift in criterion (response bias).

Why is this distinction relevant? So far research has conceptualized and tested accuracy in different ways creating a debate concerning what accuracy means (Miller, 2018). Understanding whether LGB individuals are more accurate than heterosexual people because of a clearer differentiation or because of a shift in criterion will contribute to the understanding of how gaydar judgments are made and what accuracy implies. LGB individuals may be better at differentiating speakers correctly because they are more familiar with gay/lesbian speakers and more sensitive to cues indicative of SO (Brambilla et al., 2013). In fact, LGB individuals are not only able to easily detect signals of stigmatization that define their social status and group membership (Abrams & Giles, 2004), but they also pay more attention to details than heterosexual people as they engage in a more analytic than global perceptual style (Colzato, Van Hooidonk, Van Den Wildenberg, Harinck, & Hommel, 2010). This process can be socially rewarding and motivated by the need to recognize other gay men and lesbian women who are part of their community (Colzato et al., 2010; Shelp, 2003). A shift in the criterion with respect to heterosexual participants may instead reflect the assumption that some speakers are gay/lesbian and the lower fear in using gay/lesbian labels to categorize others (see Alt et al., 2020), whereas heterosexual participants automatically assume that most speakers are heterosexual (straight categorization bias).

Overview

This research examined whether LGB individuals are better at auditory gaydar than heterosexual individuals. The limited studies on this issue have mostly focused on gaydar judgments where multiple cues or only visual cues were

available. Here, we focused on voice since it has been found to be more important than visual cues in some contexts, including spontaneous categorization (Rakić et al., 2011). In two studies we tested whether LGB listeners are more accurate than heterosexual listeners when categorizing male and female speakers' SO and, if so, whether this is driven by better differentiation or a shift in criterion. We assessed gaydar accuracy in two ways. In Study 1, participants were asked to categorize the speakers in a binary choice (gay/lesbian vs. straight). In Study 2, SO judgments were provided on a Kinsey scale from "exclusively heterosexual" to "exclusively gay/lesbian." The use of different measures to assess perceived SO is particularly important in the current debate on conceptualization and operationalization of SO as a binary rather than a more fluid concept (see Miller, 2018; Painter, Fasoli, & Sulpizio, 2021) and to better understand the underlying mechanisms, namely differentiation and shift in criterion. Both studies were approved by the Psychology Ethics Committee of the University of Padova.

We hypothesized that LGB participants would be more accurate than heterosexual participants in identifying gay/lesbian speakers (Hypothesis 1). Two contrasting hypotheses were formulated regarding the reasons for their relatively better performance. Compared to heterosexual participants, we expected LGB participants to either be better at discriminating between the two SO categories of speakers (Hypothesis 2a) or to apply a shift in criterion, by considering a higher overall likelihood that speakers would be gay (reduced straight categorization bias, Hypothesis 2b) leading to higher accuracy. In Study 2, we also assessed gender typicality since it is linked to perceived SO and affects the straight categorization bias (see Lick & Johnson, 2016). We expected gay/lesbian speakers to be perceived as less gender typical than heterosexual speakers (Hypothesis 3).

Study 1

Method

Participants

One hundred and thirty-two Italian participants completed the study. After excluding those who did not report their SO ($n = 5$), the final sample consisted of 127 participants ($M = 27.81$, $SD = 9.86$) including 70 heterosexual people (33 males and 37 females), 57 LBG individuals (28 gay and bisexual men, 29 lesbian and bisexual women). Most of the participants had a university degree (50.8%, $n = 64$), and were geographically well-distributed across Italy (see Supplementary Materials for all demographics). A G*Power sensitivity power analysis with $\alpha = .05$, $1\text{-}\beta = .80$, $N = 127$, and Cohen's $f = .158$ suggested that our sample had enough power to detect a small effect size in a mixed-design (Cohen, 1988).

Speakers

Ten males and 10 females selected from a pool of voice samples previously used by Sulpizio et al. (2020, 2015). Half of the speakers self-identified as gay/lesbian and half as heterosexual, they were all young adults (age range 24–40) and Italians from the North of Italy. Vocal stimuli consisted of two neutral sentences (i.e., "The dog runs in the park"; "The English course begins on Monday") uttered by each speaker (6–8 seconds in total).

Procedure and materials

Participants were recruited through students' contacts and invited to take part in an online study about voice-based SO recognition. After consenting to take part in the study, participants listened to either 10 male or 10 female speakers' voices (half gay/lesbian and half heterosexual) presented in a randomized order. Participants were asked to indicate the speakers' SO by making a dichotomous choice (gay/lesbian vs heterosexual). After categorizing all the speakers, participants estimated the number of gay/lesbian speakers they had listened to by choosing a number between 0 and 10. Then, they reported their demographics (i.e., age, gender, sexual orientation, geographical background, and level of education) and completed an Italian adaptation of the Attitudes Toward Lesbian and Gay men scale (ATLG, Herek, 1998) by indicating their agreement with 10 items (e.g., homosexuality is a perversion) on a scale from 1 (*completely disagree*) to 7 (*completely agree*).

Results

Preliminary analyses

ATGL showed adequate reliability ($\alpha = .65$) and thus participants' ratings were averaged. The lower the score, the more positive the attitudes toward gay men and lesbian women. Overall, participants showed positive attitudes ($M = 1.80$, $SD = .84$; t-test against the midpoint of the scale: $t(126) = -29.42$, $p < .001$), but these were more positive for LGB ($M = 1.44$, $SD = .48$) than heterosexual participants ($M = 2.09$, $SD = .95$; $t(125) = -4.67$, $p < .001$).

Analyses including participants' gender did not show any significant interaction and showed the same pattern of results of the main analyses reported below. Hence, this variable was not considered any further.

Perceived SO—dichotomous choice

We examined the data in two ways as in seminal gaydar work (Rule & Ambady, 2008). First, we looked at the descriptive accuracy for each target group. For each participant and target group we calculated the percentage of correct categorization by dividing the correct answers (multiplied by 100) by the total number of speakers. We then submitted these scores to a 2 (Speaker SO: gay/lesbian vs. heterosexual) X 2 (Speaker gender: male vs. female) X 2

(Participant SO: LGB vs. heterosexual) repeated measures ANOVA where the first variable was a within-participants factor and the others were between-participants factors. A main effect of Speaker gender, $F(1,123) = 6.68$, $p = .011$, $h^2 = .05$, indicated that on average female speakers ($M = 58.9\%$, $SD = 14.60$) were more likely to be categorized correctly than male speakers ($M = 51.87\%$, $SD = 15.72$). Also, a significant main effect of Speaker SO, $F(1,123) = 80.46$, $p < .001$, $\eta_p^2 = .39$, showed that heterosexual speakers ($M = 70.23\%$, $SD = 22.37$) were more likely to be correctly categorized than gay/lesbian speakers ($M = 40.47\%$, $SD = 25.51$). This effect was qualified by a significant interaction with Participant SO, $F(1,123) = 4.20$, $p = .042$, $\eta_p^2 = .04$. Pairwise comparisons (Bonferroni correction) showed that LGB participants ($M = 45.61\%$, $SD = 26.39$) were better than heterosexual participants ($M = 36.28\%$, $SD = 24.15$; $p = .03$) at categorizing gay/lesbian speakers, whereas no difference was found for heterosexual speakers ($M_{LGB} = 68.07\%$, $SD = 21.22$ and $M_{heterosexual} = 72\%$, $SD = 23.19$; $p = .25$). No other significant effects or interactions were found (Fs < 1.03, ps $> .31$).

These differences should be interpreted with caution. If accuracy is defined in absolute terms, scores should exceed chance level. Hence, we performed one-sample t-tests against 50%. These analyses showed that heterosexual speakers were overall correctly categorized above chance level, $t(126) = 10.19$, $p < .001$, whereas gay/lesbian speakers were categorized below chance by participants, $t(126) = -4.21$, $p < .001$. Analyses considering LGB and heterosexual participants separately showed that both participant groups categorized heterosexual speakers above chance (ts > 6.39, ps $< .001$), while gay/lesbian speakers were categorized around chance level by LGB participants, $t(126) = -1.25$, $p = .215$, and below chance by heterosexual participants, $t(126) = -4.75$, $p < .001$. Thus, an above chance accuracy emerged for heterosexual speakers, whereas accuracy for gay/lesbian speakers never exceeded the chance level indicating that both participant groups showed a straight categorization bias, but LGB participants did so to a lower degree.

Second, to identify the potential role of the criterion shift (response bias) and better understand the straight categorization bias, we performed a signal detection analysis (Stanislaw & Todorov, 1999). We calculated *hit* rates and *false alarm* rates for each participant and for gay and heterosexual speakers. We considered as *hit* rate the number of correct categorizations for gay/lesbian speakers and as *false alarms* the number of categorization of heterosexual speakers as gay/lesbian. Hit rates and false alarms were divided for the number of gay/lesbian and heterosexual speakers, respectively. We then calculated *d'* that is considered a measure of accuracy or correct discriminability, and *c* as a measure of response bias or shift in criterion. Specifically, positive values for *c* indicate a bias in categorizing individuals as heterosexual. The two indexes, *d'* and *c*, were submitted to a 2 (Speaker gender: male vs. female) X 2 (Participants SO: LGB vs. heterosexual) univariate ANOVA. Analysis on *d'*

showed only a significant main effect of Speaker gender, $F(1,123) = 5.55$, $p = .020$, $\eta_p^2 = .04$, indicating that accuracy was higher for female ($M = .81$, $SD = 1.43$) than male speakers ($M = .20$, $SD = 1.44$). Analysis on c showed instead a significant main effect of participant SO, $F(1,123) = 4.31$, $p = .040$, $\eta_p^2 = .04$, such that LGB participants ($M = .48$, $SD = .94$) were less likely to engage in the straight categorization bias than heterosexual participants ($M = .85$, $SD = 1.02$), supporting Hypothesis 2b rather than Hypothesis 2a.

Estimated number of gay/lesbian speakers

A 2 (Speaker gender: male vs. female) X 2 (Participant SO: LGB vs. heterosexual) univariate ANOVA was performed. Only a main effect of participant SO, $F(1,123) = 4.76$, $p = .031$, $\eta_p^2 = .04$, emerged. LGB participants ($M = 3.67$, $SD = 1.91$) estimated a higher number of gay/lesbian speakers in the speaker sample than heterosexual participants ($M = 3.01$, $SD = 1.47$). However, both groups of participants underestimated the number of gay/lesbian speakers they listened to as shown by t-tests against the real number of gay/lesbian speakers (5), ts < -5.27, ps $< .001$. Hence, these findings provide further indirect support for Hypothesis 2b.

Discussion

Study 1 showed that both heterosexual and LGB participants were accurate in judging heterosexual speakers but were largely inaccurate when categorizing gay/lesbian speakers. In line with Hypothesis 1, LGB participants outperformed their heterosexual counterparts in identifying gay/lesbian speakers, but their performance did not exceed chance. The relative difference between LGB and heterosexual participants are likely to reflect a differential degree of the straight categorization bias (Lick & Johnson, 2016). In line with Hypothesis 2b, the response bias was stronger for heterosexual than for LGB participants suggesting a shift in criterion rather than an accurate differentiation. This difference is evident not only in the main dependent variable (SO binary choice), but also in the estimate of the total number of gay/lesbian speakers among the voice sample. Moreover, we found that accuracy to be higher for female than male speakers both in terms of percentage of correct classification and d'. This result seems in line with previous findings observed for face-based gaydar (Brewer & Lyons, 2016; Lyons et al., 2014) and for gaydar based on nonverbal behavior (Ambady et al., 1999). Although informative, these findings only assessed perceived SO on a binary choice. Recent research (Painter et al., 2021) has shown that auditory gaydar accuracy varies when SO is assessed on a Kinsey-like scale that presents a continuum from exclusively heterosexual and exclusively gay/lesbian rather than a binary choice. We addressed this issue in Study 2.

Study 2

Study 2 aimed to conceptually replicate and extend results of Study 1 in several respects. As pointed out by Miller (2018), in the gaydar literature, SO is very often considered a binary concept and participants are forced to choose between the gay/lesbian or heterosexual categories. However, SO is a more fluid concept (Diamond, 2016). Also, the dichotomous forced-choice measure used in Study 1 may not be the best way to test people's capacity to differentiate between gay/lesbian and heterosexual speakers predicted in Hypothesis 2a. Interval scales may be better suited as they are more sensitive to relative differences in ratings (see Figure 1 for an example). Hence, in Study 2 we asked participants to rate speakers' SO on a Kinsey-like scale. Moreover, to generalize our findings, we considered a different and larger number of male and female speakers. Hypotheses 1 and 2a/2b were the same as in Study 1. Moreover, we tested here the prediction that the gay/lesbian speakers would be perceived as less gender typical than heterosexual speakers (Hypothesis 3).

Method

Participants

After excluding participants who did not complete the study (n = 36), the final sample consisted of 192 Italian participants (M_{age} = 34.70, SD = 11.48). They were 82 heterosexual (40 males and 42 females) and 110 LGB (38 gay men, 63 lesbian women and 9 bisexual women) participants. The majority was employed (69.8%, n = 134), from the North of Italy (87.5%, n = 168), Catholic (48.4%, n = 92), and preferred not to report their political orientaiton (39.5%, n = 75; see Supplementary Materials for all demographics). A G*Power sensitivity power analysis with $\alpha = .05$, $1\text{-}\beta = .80$, N = 192, and Cohen's f = .12 suggested that our sample had enough power to detect a small effect size in a mixed-design (Cohen, 1988).

Speakers

Speakers' recruitment and recording were conducted by psychology students as part of a research activity. Only recordings of good quality (no background noise, clear voice and adequate volume) were retained. All speakers uttered the same neutral sentences (i.e., "the housing market has changed in the last years and prices for a one-bedroom apartments are still high. After a long house hunting, I found a place") lasting 14–16 seconds. Recordings of 16 male speakers (8 self-identified as gay and 8 self-identified as heterosexual) and 15 female speakers (8 self-identified as lesbian and 7 as heterosexual) were used in the study. They were all young (age range 19–36) and from the north of Italy.

Procedure and measures

Listener participants were recruited online by the psychology students. After consenting to take part in the study, they listened to either the male or female speakers, one after the other in counterbalanced orders. After listening to each speaker, participants rated his/her SO on a scale from 1 (*exclusively heterosexual*) to 7 (*exclusively gay/lesbian*). Also, they rated the speaker's femininity, masculinity, and spontaneity on a scale from 1 (*not at all*) to 7 (*completely*). For results on spontaneity see the Supplementary Materials. Next, participants completed the ATLG scale as in Study 1 (α = .88) and reported their demographic information (e.g., gender, age, sexual orientation). Participants also estimated how many of the speakers they listened to were gay/lesbian (0–16/15) and reported whether they knew any of them. Finally, they reported their gaydar beliefs by indicating how difficult it was to perform the gaydar task and how easy they believed it is to detect gay and lesbian speakers. Answers were provided on a scale from 1 (*not at all*) to 7 (*completely*). At the end they were thanked and debriefed.

Results

Preliminary analyses

As in Study 1, participants ATLG ratings were averaged so that the lower the score, the more positive the attitudes. Overall participants reported positive attitudes, and LGB participants (M = 1.36, SD = .31) reported more positive attitudes than heterosexual participants (M = 2.61, SD = 1.18), $t(189) = 10.68$, $p < .001$. Moreover, LGB participants (M = 44.36, SD = 58.28) reported to know more gay and lesbian people than heterosexual participants (M = 3.49, SD = 5.63) did, $t(190) = 6.35$, $p < .001$.

Analyses also showed that participants found the task more difficult when they had to rate female (M = 5.49, SD = 1.53) than male speakers (M = 4.45, SD = 1.6), $t(190) = 4.52$, $p < .001$, and heterosexual participants (M = 5.43, SD = 1.55) found the task more difficult than LGB participants (M = 4.57, SD = 1.66), $t(190) = 3.63$, $p < .001$. Moreover, participants believed that it is easier to detect SO of gay (M = 4.00, SD = 1.74) than lesbian speakers (M = 2.86, SD = 1.86), $t(184) = 6.80$, $p < .001$. However, whereas heterosexual participants (M = 2.33, SD = 1.96) believed it is more difficult to detect lesbian speakers than LGB participants (M = 3.24, SD = 1.70, $t(185) = 3.37$, $p = .001$) did, both groups had similar beliefs about how easy it is to detect gay men's SO ($M_{heterosexual}$ = 4.00, SD = 1.80 vs M_{LGB} = 4.06, SD = 1.71, $t(188) = -.22$, $p = .83$).

Analyzes including participants' gender did not show any interactions with other variables that never reached conventional levels of significance. Hence, this variable was not further considered.

Perceived SO—Kinsey-like scale

A 2 (Speaker SO: gay/lesbian vs. heterosexual) X 2 (Speaker gender: male vs. female) X 2 (Participant SO: LGB vs. heterosexual) repeated measures ANOVA, where the first variable was a within-participants factor and all the others were between-participants variables, was performed on participants' ratings. A significant main effect of speaker SO, $F(1,187) = 68.60$, $p < .001$, $\eta_p^2 = .27$, and of speaker gender, $F(1,187) = 4.05$, $p = .046$, $\eta_p^2 = .02$, emerged. These effects were qualified by a significant interaction between the two variables, $F(1,187) = 54.72$, $p < .001$, $\eta_p^2 = .23$. Pairwise comparisons (Bonferroni correction) indicated that participants reported higher ratings for gay ($M = 4.24$, $SD = .83$) than heterosexual male speakers ($M = 3.08$, $SD = .94$; $p < .001$), whereas no difference emerged for female speakers ($M_{lesbian} = 3.46$, $SD = .98$ and $M_{heterosexual} = 3.41$, $SD = 1.08$; $p = .54$).

Also, LGB participants ($M = 3.79$, $SD = .75$) reported overall higher ratings than heterosexual participants ($M = 3.24$, $SD = .80$), $F(1,187) = 24.25$, $p < .001$, $\eta_p^2 = .11$, thus showing a reduced straight-categorization bias, in line with Hypothesis 2b. A significant interaction between speaker SO, speaker gender and participant SO, $F(1,187) = 4.27$, $p = .04$, $\eta_p^2 = .02$, showed that the effect illustrated above occurred specifically for gay speakers. Indeed, pairwise comparisons (Bonferroni correction) showed that LGB participants ($M = 4.47$, $SD = .83$) perceived gay male speakers as more gay than heterosexual participants ($M = 3.92$, $SD = .72$; $p = .001$), whereas no difference emerged on ratings for the heterosexual male speakers ($M_{LGB} = 3.17$, $SD = .89$ and $M_{heterosexual} = 2.94$, $SD = 1.00$; $p = .26$). On the contrary, LGB participants perceived both lesbian ($M = 3.72$, $SD = .86$) and heterosexual ($M = 3.80$, $SD = .96$) female speakers as more lesbian than heterosexual participants ($M_{lesbian\_speakers} = 3.14$, $SD = 1.04$ and $M_{heterosexual\_speakers} = 2.93$, $SD = 1.04$; $ps < .01$). No other significant main effects or interactions emerged ($Fs < 2.21$, $ps > .14$).

This analysis suggests that participants distinguished gay and heterosexual male speakers in relative terms. However, to prove accuracy in absolute terms, means for heterosexual speakers should be below and means for gay speakers above the scale midpoint (4). Thus, we performed one sample *t*-tests against the scale midpoint for each group of speakers, separately for gay/lesbian and heterosexual speakers. With regards to male speakers, both LGB and heterosexual participants correctly rated the heterosexual speakers below the scale midpoint ($ts < -6.88$, $ps > .001$), while gay speakers were rated around the scale midpoint by heterosexual participants, $t(41) = -.70$, $p = .48$, but were correctly identified by LGB participants since the mean was significantly above the scale midpoint, $t(59) = 4.35$, $p < .001$. With regards to female speakers, heterosexual participants correctly rated heterosexual speakers below the midpoint, $t(39) = -6.50$, $p < .001$, whereas LGB participants rated these speakers around the scale

midpoint, $t(49) = -1.52$, $p = .13$. At the same time, both heterosexual and LGB participants rated lesbian speakers below the scale midpoint ($ts < -2.27$, $ps < .03$).

Differentiation or shift in criterion?

A different way to assess the degree of differentiation is to analyze the difference score between the two judgments, by subtracting the rating of the heterosexual from that of the gay/lesbian speakers. Greater scores indicate a greater differentiation between the two categories. Difference scores were submitted to a 2 (Speaker gender: male vs. female) X 2 (Participant SO: LGB vs. heterosexual) mixed ANOVA. A significant main effect for Speaker gender, $F(1,187) = 54.72$, $p < .001$,$h^2 = .23$, indicated a better differentiation between gay/lesbian and heterosexual speakers for male ($M = 1.16$, $SD = 1.03$) than for female speakers ($M = .05$, $SD = .95$). Also, a significant but weaker interaction with Participant SO, $F(1,187) = 4.27$, $p = .04$,$h^2 = .02$, emerged. For male speakers, LGB participants ($M = 1.29$, $SD = .89$) slightly outperformed heterosexual participants ($M = .98$, $SD = 1.29$). For female speakers, heterosexual participants ($M = .21$, $SD = 1.02$) slightly outperformed LGB participants ($M = -.08$, $SD = .88$). However, neither of these trends was significant ($ps > .13$). Together, these findings suggest that LGB members are no better than heterosexual participants at differentiating the SO of gay/lesbian and heterosexual speakers.

The results reported above considered perceived SO measured on a Kinsey-like scale supposedly allowing for a conceptualization of SO as a continuum. However, people often think about SO as a binary. In order to allow a direct comparison between the two studies and further test differentiation and shift in criterion, a signal detection analysis was performed. To do that, we converted the responses into dichotomous responses. Depending on whether responses were located on one or the other side of the scale they were recoded as "gay/lesbian" if ratings were above and "heterosexual" if they were below 4. Although this re-coding strategy has been recently criticized because it reiterates a SO binary conceptualization (Miller, 2018), it is useful to understand how listeners make judgments and to compare accuracy as usually conceptualized in gaydar research. This recoding allowed us to calculate hits and false alarms, and then compute the *d'* (discrimination) and *c* (criterion or response bias) scores as in Study 1. Positive values for *c* indicate a bias in categorizing individuals as heterosexual. The two indexes, *d'* and *c*, were submitted to a 2 (Speaker gender: male vs. female) X 2 (Participants SO: LGB vs. heterosexual) ANOVA. Analysis on *d'* showed only a significant main effect of Speaker gender, $F(1,188) = 26.40$, $p < .001$, $\eta_p^2 = .12$, indicating that accuracy was higher for male ($M = .47$, $SD = .96$) than for female speakers ($M = -.39$, $SD = 1.37$). No other significant effects or interactions were found ($Fs < 2.01$, $ps >$

.16). As in Study 1, there was no significant main effect or interactions involving participant SO, suggesting that LGB participants were no better at differentiating the two groups than were their heterosexual counterparts.

Analysis on *c* scores showed instead a significant main effect of participant SO, $F(1,188) = 9.54$, $p = .002$, $\eta_p^2 = .05$, such that LGB participants ($M = -.28$, $SD = .71$) were less likely to engage in the straight categorization bias than heterosexual participants ($M = .07$, $SD = .91$), thereby replicating the result pattern of Study 1 and supporting Hypothesis 2b. A significant main effect for Speaker Gender, $F(1,188) = 6.96$, $p = .009$, $\eta_p^2 = .04$, also revealed that the response bias was greater for female ($M = -.02$, $SD = .99$) than for male speakers ($M = -.27$, $SD = .59$). The interaction between the two variables was not significant, $F(1,188) = 2.02$, $p = .16$, $\eta_p^2 = .01$.

Estimated number of gay/lesbian speakers

A 2 (Speaker gender: male vs. female) X 2 (Participant SO: LGB vs. heterosexual) univariate ANOVA was performed on the number of gay/lesbian speakers participants believed to have recognized. Only a significant main effect of participant SO, $F(1,172) = 52.33$, $p < .001$,$h^2 = .23$, emerged. LGB participants ($M = 7.09$, $SD = 2.73$) reported to have recognized more gay/lesbian speakers than heterosexual participants ($M = 4.36$, $SD = 2.11$), attesting to a lower straight categorization bias among LGB participants and supporting Hypothesis 2b. Also, *t*-test against the actual number of gay/lesbian speakers (8) showed that both LGB and heterosexual participants underestimated the number of gay/lesbian speakers they listened to, $ts < -3.35$, $ps < .001$. No other significant effects or interactions were found ($Fs < 1.26$, $ps > .26$).

Gender typicality

Masculinity and femininity ratings were negatively correlated for all group of speakers ($rs < -.92$, $ps < .001$). A gender typicality score was calculated by subtracting ratings on femininity from those on masculinity for male speakers and vice versa for female speakers. Hence, the higher the score, the more gender typical the speaker was perceived. This score was submitted to a 2 (Speaker SO: gay/lesbian vs. heterosexual) X 2 (Speaker gender: male vs. female) X 2 (Participant SO: LGB vs. heterosexual) repeated measures ANOVA, the first as within-participants and the others as between-participants variables. Significant main effects of speaker SO, $F(1,186) = 69.50$, $p < .001$,$h^2 = .27$, and of speaker gender, $F(1,186) = 5.60$, $p = .02$,$h^2 = .03$, were qualified by a significant interaction between these two factors, $F(1,186) = 93.58$, $p < .001$,$h^2 = .33$. Pairwise comparisons (Bonferroni correction) showed that heterosexual male speakers ($M = 2.86$, $SD = 1.86$) were perceived as more gender typical than gay male speakers ($M = .92$, $SD = 1.94$; $p < .001$). On the

contrary, no difference emerged for the perception of female speakers' gender typicality ($M_{lesbian}$ = 2.56, SD = 1.96 vs $M_{heterosexual}$ = 2.42, SD = 2.09; p = .36). No other significant effects or interactions were found (Fs < 2.80, ps > .10).

Correlations

We analyzed correlations between perceived SO and gender typicality ratings separately for speaker SO and gender. Gender typicality was negatively correlated with perceived SO for both gay and heterosexual male speakers (rs > −.58, ps < .001) and for both lesbian and heterosexual female speakers (rs > −.50, ps < .001).

Moreover, correlations between d' and c and gender typicality for male and female speakers were performed. Analyses only showed that the response bias (c) significantly and positively correlated with gender typicality for both male, $r(101)$ = .22, p = .03, and female speakers, $r(90)$ = .36, p = .001. Hence, the more the speakers were perceived as gender typical the more participants engaged in the straight categorization bias. However, this was true for heterosexual (male speakers: $r(41)$ = .31, p = .049 and female speakers; $r(40)$ = .53, p < .001) but not for LGB participants (rs < .16, ps > .21), suggesting that heterosexual participants were more likely to consider gender typicality when making gaydar judgments than LGB participants did.

Finally, we tested the correlations between the number of gay and lesbian people participants knew and ATGL with measures of accuracy, namely differentiation, d prime and c, separately for speaker gender and participants' SO. None of these correlations was significant for participants' contacts with gay and lesbian people (rs < .25, ps > .08). ATGL correlated positively with c when male speakers and LGB participants were considered, $r(60)$ = .28, p = .03. Hence, the more they held positive attitudes, the higher was the likelihood to label male speakers as gay. No other significant correlation was found (rs < .19, ps > .20).

Discussion

Study 2 showed a difference in relative terms that partially confirmed Hypothesis 1. Whereas heterosexual participants' ratings did not exceed the scale midpoint indicating overall inaccuracy as in previous studies (Sulpizio et al., 2020, 2015), LGB participants rated gay, but not lesbian speakers, on the correct side of the scale. Importantly, when looking at the differentiation in SO ratings for gay/lesbian and heterosexual speakers, LGB participants tended to outperform heterosexual participants when rating male speakers, but the opposite pattern occurred when participants were rating female speakers. Hence, LGB did not seem to be better at differentiating speakers' SO.

People, especially heterosexual people, may have a binary conception of SO and may think in these terms even when a Kinsey-like scale is provided. Recoding participants' ratings in a binary category allowed us to compare results across the two studies. Analyses on recorded ratings demonstrated that LGB participants showed the same criterion shift (response bias) observed in Study 1, but not higher accuracy, confirming Hypothesis 2b. Also, providing further support to this hypothesis, LGB participants estimated to have listened to a higher number of gay/lesbian speakers than heterosexual participants did.

Contrary to Study 1, we found here that accuracy was higher for male than female speakers. Indeed, in Study 2 we did not observe a relative difference between lesbian and heterosexual female speakers. This result may be due to the different SO measures and/or voice samples used across the two studies. At the same time, participants reported that they found the SO categorization less difficult when it concerned male than female voices and believed generally that it is easier to detect gay men than lesbian women from voice. This result is in line with work showing that individuals endorse essentialist beliefs that gay and heterosexual men, more than heterosexual and lesbian women's voices, are distinct making it easier for listeners to detect male speakers' SO (Fasoli, Hegarty, & Frost, 2021). Hence, it may be that our participants were hesitant to label female speakers as lesbian because they believed that voice is unlikely to convey information about women's SO.

Moreover, gay male speakers were perceived as less gender typical, namely less masculine, than heterosexual male speakers, partially confirming Hypothesis 3. No difference on gender typicality was found for female speakers. Also, these perceptions were not moderated by participants' SO. In line with the straight categorization bias (Lick & Johnson, 2016), gender typicality was associated with perceived SO of the speakers and response bias indicating that perceiving speakers as gender typical led participants to assume they were straight. Interestingly, no correlation between the number of gay/lesbian people participants knew and accuracy or response bias emerged, suggesting that familiarity may not relate to higher accuracy or response bias when voice is concerned. This speaks against the idea that LGB individuals are better at auditory gaydar because they are exposed to a greater variety of gay/lesbian voices than heterosexual people, a prediction that assumes actual differences in gay/lesbian and heterosexual voices and that have been questioned (see Miller, 2018). Instead, it appears that what matters is the less stigmatization attached to labeling others as gay (see Alt et al., 2020). Indeed, the more positive attitudes toward gay men were associated with LGB speakers' engagement in criterion shift, namely being less restrictive in categorizing others as gay. However, this was true only for male speakers.

General discussion

This research examined whether LGB individuals have a better auditory gaydar than heterosexual people. We consistently found that LGB participants did not exceed in accuracy but rather engaged in a criterion shift that makes them less likely to show a straight categorization bias (Lick & Johnson, 2016) and more likely to label others as gay/lesbian. Although we found some relative differences in SO categorization, overall, we observed an auditory gaydar inaccuracy in line with previous research (see Sulpizio et al., 2015, 2020; for similar results on Italian speakers). Percentages of correct SO categorization for gay and lesbian speakers never exceeded chance in Study 1 and only in one case ratings for gay, but not lesbian speakers, were on the correct side of the Kinsey-like scale in Study 2. Similarly, participants always underestimated the number of gay/lesbian speakers they listened to.

This research contributes to the literature in several respects. First, our findings support overall auditory gaydar inaccuracy, even when LGB participants were involved (Rieger et al., 2010; Valentova & Havlíček, 2013). Second, we did not find any support for the idea that gaydar is an "innate" or "adaptive" skill that only LGB individuals have (Shelp, 2003; Woolery, 2007). Instead, we found that LGB participants were less "rigid" in their SO categorization and more likely to categorize speakers as gay as shown by their weaker straight categorization bias. As it happens for heterosexual people, this response bias may be driven by the reluctance to label others as gay/lesbian (see Brewer & Lyons, 2016) possibly because this is considered as stigmatizing (Alt et al., 2020). LGB individuals may not see labeling others as gay/lesbian as harmful and hence be more likely to categorize others as such.

Alternatively, it could be that participants of different SO have different subjective baselines concerning the likelihood of someone being gay/lesbian. Lick and Johnson (2016) showed that the straight categorization bias in samples that were mostly heterosexuals was not triggered by base rate information but rather than by gendered cues in the stimuli. Study 2 showed a correlation between gender typicality and response bias for heterosexual but not for LGB participants. This may imply that LGB individuals do not rely on gender typicality when making auditory gaydar judgments as heterosexual individuals do. This is an interesting possibility aligning with the suggestion by Colzato et al. (2010) concerning different attentional focus of heterosexual and LGB people that needs further investigation.

This is one of the few studies to examine auditory gaydar across LGB and heterosexual listeners and to compare gaydar judgments for both male and female speakers. Our findings showed a higher accuracy for female speakers in Study 1, but for male speakers in Study 2. Literature has provided mixed results with regards with the gender of the target. Studies on visual gaydar have shown better accuracy for female than male targets (Ambady et al., 199;

Brewer & Lyons, 2016) that would be in line with results of Study 1. Studies on voices have found that judgments were overall inaccurate for both male and female speakers' SO (Kachel et al., 2018, Kachel et al., 2020; Sulpizio et al., 2020, 2015) and that the straight categorization bias based on averaged voices was stronger for male than for female (Kachel et al., 2018). In other research, accuracy depended on both listeners and speakers' gender and SO (e.g., lesbian listeners' accuracy was worse when judging male speakers from speech, Rieger et al., 2010).

One could look for explanations in differences across studies. The voice samples and recordings were different, with Study 2 involving more speakers and longer audio stimuli, presumably leading to more reliable findings. Smaller voice samples imply less intra-group variation (see Kachel et al., 2018) and less representative samples. Participants also reported that it was particularly difficult to categorize female speaker SO in Study 2. This may be due to the type of measure (Kinsey-like scale) used in Study 2. This type of scale assumes that SO is fluid, whereas people often think of SO as a binary concept. Sexual fluidity seems to be more common among women than men (see Diamond, 2016), possibly explaining the difficulty in providing clear categorization for female targets. Moreover, participants believed it is more difficult to detect lesbian women than gay men, a result in line with the idea that voice reflects essential and immutable differences between LG and heterosexual people (Fasoli et al., 2021). If participants in the two studies had different gaydar beliefs, such difference may explain contrasting results. Future studies need to address the impact of beliefs, type of stimuli and measure on gaydar judgments (see Painter et al., 2021).

These findings are particularly important with regards to the consequences of auditory gaydar. Believing that voice is a SO cue and that auditory gaydar exists is associated with stigmatization of gay/lesbian-sounding speakers by heterosexual people (Fasoli et al., 2021). Moreover, speakers perceived to be gay are at risk of stereotyping, social distance and discrimination (Fasoli & Maass, 2018). Knowing how gaydar judgments are made, allow scholars to design interventions aimed to deconstruct gaydar beliefs and prevent voice-based stigmatization perpetrated by heterosexual people. At the same time, knowing that LGB people are inaccurate in their gaydar judgments suggests that they are also at risk of miscategorizing individuals with all the consequences that this implies for social interactions (e.g., forming a community, difficulties in dating, etc.). Still, LGB people's shift in criterion also testifies a change in beliefs that may lead to deconstruct rigid views of heterosexuality supporting heteronormativity.

Limitations and future directions

This research involves some limitations. First of all, findings are specific to two different voice samples used here, which makes direct comparison and generalization difficult. Indeed, intra-speaker vocal variation is remarkable and affects the way speakers' SO is perceived (Kachel et al., 2017; Kachel et al., 2018). Moreover, our speakers and listeners were all Italian. Future studies should aim to expand this investigation by using larger voice samples and by considering speakers and listeners of different languages (Sulpizio et al., 2020, 2015). Examining different cultures may be particularly important as endorsement of gender and sexual stereotypes vary. This research was conducted in Italy, where not only heterosexual but also LG people have reported negative reactions toward other LG individuals depending on their gender typicality (see Hunt, Fasoli, Carnaghi, & Cadinu, 2016; Salvati, Pistella, Giacomantonio, & Baiocco, 2018).

The present research showed similar results on criterion shift when different measures of SO were considered. One may argue that the Kinsey-like scale simply reflects a confidence rating. However, recent work relying on both Kinsey-like scale and confidence ratings showed different patterns of results on the two measures, disconfirming this possibility (Painter et al., 2021). Still, the Kinsey-like scale may represent a proxy of speaker SO prototypicality rather than a measure assessing fluid SO. Future studies should consider alternative SO measures disentangling SO from prototypicality as well as expanding the SO choices by including other categories (e.g., bisexual, asexual, etc.).

In our studies we limited our investigation to SO categorization without considering participants' motives. Research has indicated that individuals may want to guess others' SO for mating interests (Rule, Rosen, Slepian, & Ambady, 2011) or to avoid gay/lesbian individuals who represent a "threat" for their SO identity (Plant, Zielaskowski, & Buck, 2014). Future research should consider differences in gaydar judgments among listeners of distinct SO (e.g., bisexuals, asexual) and gender identity (e.g., trans and non-binary individuals) and the motivations behind their gaydar judgments.

Finally, very few studies have compared SO judgments based on vocal and visual cues (Kachel et al., 2020; Valentova & Havlíček, 2013) and less is known of such comparison when both heterosexual and LGB individuals are involved. Similar mechanisms have been explored as possible explanations for visual and auditory gaydar (e.g., straight categorization bias). However, more research is needed to understand whether processing of visual and vocal information show similar patterns also among LGB people (Colzato et al., 2010). For instance, in our findings gender typicality appeared to matter less for criterion shifts among LGB individuals.

Conclusion

Overall, both LGB and heterosexual individuals appear to be quite inaccurate in their gaydar judgments. Nevertheless, LGB individuals consistently showed a shift in criterion that made them more likely to categorize others as gay or lesbian. This can have important implications for everyday interactions and stigmatization: Differently from heterosexual people, LGB individuals engage less in the default assumption of heterosexuality that is at the basis of sexual stigma.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Fabio Fasoli http://orcid.org/0000-0002-8043-1630

References

Abrams, J. R., & Giles, H. (2004). An intergroup to communicating stigma: Gays and lesbians. In S. H. Ng, C. N. Candlin, & C. Y. Chiu (Eds.), *Language Matters: Communication, Identity, and Culture*, (pp. 27–61). Hong Kong: City University of Hong Kong Press.

Alt, N. P., Lick, D. J., & Johnson, K. L. (2020). The straight categorization bias: A motivated and altruistic reasoning account. *Journal of Personality and Social Psychology*, *119*(6), 1266–1289. doi:10.1037/pspi0000232

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*(3), 538–547. https://psycnet.apa.org/doi/10.1037/0022-3514.77.3.538

Baker, P. (2003). *Polari-the lost language of gay men*. London: Routledge.

Barton, B. (2015). How like perceives like: Gay people on "gaydar." *Journal of Homosexuality*, *62*(12), 1615–1637. doi:10.1080/00918369.2015.1091207

Berger, G., Hank, L., Rauzi, T., & Simkins, L. (1987). Detection of sexual orientation by heterosexuals and homosexuals. *Journal of Homosexuality*, *13*(4), 83–100. doi:10.1300/J082v13n04_05

Brambilla, M., Riva, P., & Rule, N. O. (2013). Familiarity increases the accuracy of categorizing male sexual orientation. *Personality and Individual Differences*, *55*(2), 193–195. doi:10.1016/j.paid.2013.02.023

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

Chauncey, G. (1994). *Gay New York: Gender, urban culture, and the making of the gay male world, 1890-1940*. New York, NY: Basic Books.

Clarke, V., Hayfield, N., & Huxley, C. (2012). Lesbian, gay, bisexual and trans appearance and embodiment: A critical review of the psychological literature. *Psychology of Sexualites Review*, *3*, 51–70.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Colzato, L. S., Van Hooidonk, L., Van Den Wildenberg, W., Harinck, F., & Hommel, B. (2010). Sexual orientation biases attentional control: A possible gaydar mechanism. *Frontiers in Psychology*, *1*, 13. doi:10.3389/fpsyg.2010.00013

Cox, W. T., Devine, P. G., Bischmann, A. A., & Hyde, J. S. (2016). Inferences about sexual orientation: The roles of stereotypes, faces, and the gaydar myth. *The Journal of Sex Research*, *53*(2), 157–171. doi:10.1080/00224499.2015.1015714

Daniele, M., Fasoli, F., Antonio, R., Sulpizio, S., & Maass, A. (2020). in press). Gay voice: Stable marker of sexual orientation or flexible communication device? *Archives of Sexual Behavior*, *49*(7), 2585–2600. doi:10.1007/s10508-020-01771-2.

Diamond, L. M. (2016). Sexual fluidity in male and females. *Current Sexual Health Reports*, *8*(4), 249–256. doi:10.1007/s11930-016-0092-z

Fasoli, F., Hegarty, P., & Frost, D. M. (2021). Stigmatization of 'gay-sounding' voices: The role of heterosexual, lesbian, and gay individuals' essentialist beliefs. *British Journal of Social Psychology*, *60*(3), 826–850. doi:10.1111/bjso.12442

Fasoli, F., Hegarty, P., Maass, A., & Antonio, R. (2018). Who wants to sound straight? Sexual majority and minority stereotypes, beliefs and desires about auditory gaydar. *Personality and Individual Differences*, *130*, 59–64. doi:10.1016/j.paid.2018.03.046

Fasoli, F., & Maass, A. (2018). Voice and prejudice: The social costs of auditory gaydar. *Atlantic Journal of Communication*, *26*(2), 98–110. doi:10.1080/15456870.2018.1432617

Gaudio, R. P. (1994). Sounding gay: Pitch properties in the speech of gay and straight men. *American Speech*, *69*(1), 30–57. doi:10.2307/455948

Herek, G. M. (1988). Heterosexuals' attitudes toward lesbians and gay men: Correlates and gender differences. The Journal of Sex Research, *25*, 451–411. https://doi.org/10.1080/00224498809551476

Hunt, C. J., Fasoli, F., Carnaghi, A., & Cadinu, M. (2016). Masculine self-presentation and distancing from femininity in gay men: An experimental examination of the role of masculinity threat. *Psychology of Men & Masculinity*, *17*(1), 108. doi:10.1037/a0039545

Kachel, S., Radtke, A., Skuk, V. G., Zäske, R., Simpson, A. P., & Steffens, M. C. (2018). Investigating the common set of acoustic parameters in sexual orientation groups: A voice averaging approach. *PloS One*, *13*(12), e0208686. doi:10.1371/journal.pone.0208686

Kachel, S., Simpson, A. P., & Steffens, M. C. (2017). Acoustic correlates of sexual orientation and gender-role self-concept in women's speech. The Journal of the Acoustical Society of America, *141*(6), 4793–4809. https://doi.org/10.1121/1.4988684

Kachel, S., Simpson, A. P., & Steffens, M. C. (2018). "Do I sound straight?": Acoustic correlates of actual and perceived sexual orientation and masculinity/femininity in men's speech. *Journal of Speech, Language, and Hearing Research*, *61*(7), 1560–1578. doi:10.1044/2018_JSLHR-S-17-0125

Kachel, S., Steffens, M. C., Preuß, S., & Simpson, A. P. (2020). Gender (conformity) matters: Cross-dimensional and cross-modal associations in sexual orientation Perception. *Journal of Language and Social Psychology*, *39*(1), 40–66. doi:10.1177/0261927X19883902

Kreiman, J., Keating, P. A., Park, S. J., Rastifar, S., & Alwan, A. (2015). Within-and between-talker variability in voice quality in normal speaking situations. *The Journal of the Acoustical Society of America*, *137*(4), 2418. doi:10.1121/1.4920824

Lavan, N., Burton, A. M., Scott, S. K., & McGettigan, C. (2019). Flexible voices: Identity perception from variable vocal signals. *Psychonomic Bulletin & Review*, *26*(1), 90–102. doi:10.3758/s13423-018-1497-7

Lick, D. J., & Johnson, K. L. (2016). Straight until proven gay: A systematic bias toward straight categorizations in sexual orientation judgments. *Journal of Personality and Social Psychology*, *110*(6), 801–817. doi:10.1037/pspa0000052

Linville, S. E. (1998). Acoustic correlates of perceived versus actual sexual orientation in men's speech. *Folia Phoniatrica Et Logopaedica*, *50*(1), 35–48. doi:10.1159/000021447

Lyons, M., Lynch, A., Brewer, G., & Bruno, D. (2014). Detection of sexual orientation ("gaydar") by homosexual and heterosexual women. Archives of Sexual Behavior, *43*(2), 345–352. https://doi.org/10.1007/s10508-013-0144-7

Miller, A. E. (2018). Searching for gaydar: Blind spots in the study of sexual orientation perception. *Psychology & Sexuality*, *9*(3), 188–203. doi:10.1080/19419899.2018.1468353

Munson, B. (2007). The acoustic correlates of perceived masculinity, perceived femininity, and perceived sexual orientation. *Language and Speech*, *50*(1), 125–142. doi:10.1177/00238309070500010601

Nicholas, C. L. (2004). Gaydar: Eye-gaze as identity recognition among gay men and lesbians. *Sexuality and Culture*, *8*(1), 60–86. doi:10.1007/s12119-004-1006-1

Painter, D., Fasoli, F., & Sulpizio, S. (2021). The impact of stimuli length and analytic method on auditory gaydar. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2021.08.016

Plant, E. A., Zielaskowski, K., & Buck, D. M. (2014). Mating motives and concerns about being misidentified as gay or lesbian: Implications for the avoidance and derogation of sexual minorities. *Personality and Social Psychology Bulletin*, *40*(5), 633–645. doi:10.1177/0146167214521467

Podesva, R. J. (2007). Phonation type as a stylistic variable: The use of falsetto in constructing a persona 1. *Journal of Sociolinguistics*, *11*(4), 478–504. doi:10.1111/j.1467-9841.2007.00334.x

Rakić, T., Steffens, M. C., & Mummendey, A. (2011). Blinded by the accent! The minor role of looks in ethnic categorization. *Journal of Personality and Social Psychology*, *100*(1), 16–29. doi:10.1037/a0021522

Ravenhill, J., & de Visser, R. O. (2017). Perceptions of gay men's masculinity are associated with their sexual label, voice quality and physique. *Psychology & Sexuality*, *8*(3), 208–222. doi:10.1080/19419899.2017.1343746

Rieger, G., Linsenmeier, J. A., Gygax, L., Garcia, S., & Bailey, J. M. (2010). Dissecting "gaydar": Accuracy and the role of masculinity–femininity. *Archives of Sexual Behavior*, *39*(1), 124–140. doi:10.1007/s10508-008-9405-2

Ruben, M. A., Hill, K. M., & Hall, J. A. (2014). How women's sexual orientation guides accuracy of interpersonal judgements of other women. *Cognition and Emotion*, *28*(8), 1512–1521. doi:10.1080/02699931.2014.890093

Rule, N. O., & Ambady, N. (2008). Brief exposures: Male sexual orientation is accurately perceived at 50 ms. *Journal of Experimental Social Psychology*, *44*(4), 1100–1105. doi:10.1016/j.jesp.2007.12.001

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*(7), 881–886. doi:10.1177/0956797611412394

Rule, N. O. (2017). Perceptions of sexual orientation from minimal cues. *Archives of Sexual Behavior*, *46*(1), 129–139. doi:10.1007/s10508-016-0779-2

Salvati, M., Pistella, J., Giacomantonio, M., & Baiocco, R. (2018). Lesbians' negative affect toward sexual minority people with stereotypical masculine and feminine characteristics. *International Journal of Sexual Health*, *30*(2), 162–176. doi:10.1080/19317611.2018.1472705

Shelp, S. G. (2003). Gaydar. *Journal of Homosexuality*, *44*(1), 1–14. doi:10.1300/J082v44n01_01

Smyth, R., Jacobs, G., & Rogers, H. (2003). Male voices and perceived sexual orientation: An experimental and theoretical approach. *Language in Society*, *32*(3), 329–350. doi:10.1017/S0047404503323024

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. doi:10.3758/BF03207704

Sulpizio, S., Fasoli, F., Antonio, R., Eyssel, F., Paladino, M. P., & Diehl, C. (2020). Auditory Gaydar : Perception of Sexual Orientation Based on Female Voice. *Language and Speech*, *63*(1), 184–206. doi:10.1177/0023830919828201

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*(7), e0128882. doi:10.1371/journal.pone.0128882

Tracy, E. C., Bainter, S. A., & Satariano, N. P. (2015). Judgments of self-identified gay and heterosexual male speakers: Which phonemes are most salient in determining sexual orientation? *Journal of Phonetics*, *52*, 13–25. doi:10.1016/j.wocn.2015.04.001

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*(12), e82417. doi:10.1371/journal.pone.0082417

Woolery, L. M. (2007). Gaydar. *Journal of Homosexuality*, *53*(3), 9–17. doi:10.1300/J082v53n03_02