



Contents lists available at ScienceDirect

Journal of Research in Personality

journal homepage: www.elsevier.com/locate/jrp

Do voices carry valid information about a speaker's personality?

Julia Stern^{a,*}, Christoph Schild^{b,c}, Benedict C. Jones^d, Lisa M. DeBruine^e, Amanda Hahn^f, David A. Puts^g, Ingo Zettler^c, Tobias L. Kordsmeyer^a, David Feinberg^h, Dan Zamfirⁱ, Lars Penke^a, Ruben C. Arslan^j

^a Department of Psychology & Leibniz ScienceCampus Primate Cognition, University of Goettingen, Gosslerstrasse 14, 37073 Goettingen, Germany

^b Department of Psychology, University of Siegen, Adolf-Reichwein-Str. 2a, 57068 Siegen, Germany

^c Department of Psychology, University of Copenhagen, Øster Farimagsgade 2A, 1353 Copenhagen, Denmark

^d School of Psychological Sciences and Health, University of Strathclyde, Glasgow, UK

^e Institute of Neuroscience & Psychology, University of Glasgow, Hillhead Street, Glasgow, Scotland, UK

^f Department of Psychology, Humboldt State University, 1 Harpst Street, Arcata, CA 95521, USA

^g Department of Anthropology & Center for Brain, Behavior and Cognition, Pennsylvania State University, University Park, PA 16802, USA

^h Department of Psychology, Neuroscience, and Behaviour, McMaster University, Hamilton, Ontario L8S 4K1, Canada

ⁱ Developmental Psychology and Education, University of Toronto, 63 St. George Street Toronto, ON M5S Canada

^j Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany

ARTICLE INFO

Article history:

Received 11 April 2019

Revised 16 July 2020

Accepted 16 March 2021

Available online xxx

Keywords:

Voice pitch

Sociosexual orientation

Dominance

Big 5 personality traits

Registered Report

ABSTRACT

Research on links between peoples' personality traits and their voices has primarily focused on other peoples' personality judgments about a target person based on a target person's vocal characteristics, particularly voice pitch. However, it remains unclear whether individual differences in voices are linked to actual individual differences in personality traits, and thus whether vocal characteristics are indeed valid cues to personality. Here, we investigate how the personality traits of the Five Factor Model of Personality, sociosexuality, and dominance are related to measured fundamental frequency (voice pitch) and formant frequencies (formant position). For this purpose, we conducted a secondary data analysis of a large sample (2217 participants) from eleven different, independent datasets with a Bayesian approach. Results suggest substantial negative relationships between voice pitch and self-reported sociosexuality, dominance and extraversion in men and women. Thus, personality might at least partly be expressed in people's voice pitch. Evidence for an association between formant frequencies and self-reported personality traits is not compelling but remains uncertain. We discuss potential underlying biological mechanisms of our effects and suggest a number of implications for future research.

© 2021 Elsevier Inc. All rights reserved.

1. Introduction

1.1. Theoretical background

When meeting new people, we make spontaneous inferences and form first impressions about a wide range of characteristics (e.g. Ambady & Rosenthal, 1992). Besides physical characteristics, such as sex (Bachorowski & Owren, 1999; Puts, Apicella, & Cárdenas, 2012), body size (Feinberg, Jones, Little, Burt, & Perrett, 2005; Xu, Lee, Wu, Liu, & Birkholz, 2013), or strength (Sell et al., 2010), we also form stable impressions about other relevant characteristics, including attitudes, intentions, values, beliefs and personality traits (Borkenau & Liebler, 1992; Borkenau, Mauer,

Riemann, Spinath, & Angleitner, 2004; McAleer, Todorov, & Belin, 2014; Scherer, 1972).

While other peoples' physical appearance might be an important cue to our social evaluations (Naumann, Vazire, Rentfrow, & Gosling, 2009), peoples' voices are another factor that influences socially relevant impressions (Borkenau & Liebler, 1992; Mileva, Tompkinson, Watt, & Burton, 2018; Zuckerman & Driver, 1989). Indeed, when visual cues are absent, e.g., when listening to the radio or hearing a voice on the telephone, people still form judgments about others, based only on acoustic information (Borkenau & Liebler, 1992; Mileva et al., 2018).

Human voices and judgments based on their sound seem to have an association with important life outcomes: Studies have reported that voice characteristics predict mate choice (for an overview see Puts, Doll, & Hill, 2014), courtship outcomes (Leongómez et al., 2014), and reproductive success (Apicella, Feinberg, & Marlowe, 2007). Even voting behavior has been

* Corresponding author.

E-mail address: julia.stern@psych.uni-goettingen.de (J. Stern).

reported to be influenced by politicians' voices, in that participants preferred to vote for candidates with a lower voice pitch (the rate of vocal fold vibrations which influences perceptions of pitch, usually equated with fundamental frequency, e.g. how high or deep a voice sounds), presumably because low pitch sounds more dominant, honest, intelligent, and attractive (Klofstad, Anderson, & Nowicki, 2015; Tigue, Borak, O'Connor, Schandl, & Feinberg, 2012). Further, CEOs with lower voice pitch oversee larger companies, receive higher compensations, and enjoy longer tenures than CEOs with higher voice pitch (Mayew, Parsons, & Venkatachalam, 2013).

Voice pitch has been associated with personality judgments in multiple studies, showing that men and women with lower voice pitch are perceived as more dominant (e.g. Borkowska & Pawlowski, 2011; Collins, 2000; Hodges-Simeon, Gaulin, & Puts, 2010; Puts, Gaulin, & Verdolini, 2006; Puts, Hodges, Cárdenas, & Gaulin, 2007). Furthermore, people with higher pitched voices have been perceived to be more nervous (Apple, Streeter, & Krauss, 1979), less agreeable (Scherer, 1978), and higher in neuroticism (Aronovitch, 1976; Scherer, 1978). Moreover, men with lower voice pitch and lower formant frequencies (defined as resonant frequencies determined by the length and shape of the vocal tract and influence perceptions of vocal timbre, an example for perceivable changes in formant frequencies without changes in pitch is raising formant frequencies when inhaling helium gas) are perceived as more attractive (e.g. Collins, 2000; Feinberg et al., 2011; Hodges-Simeon et al., 2010; Jünger, Kordsmeyer, Gerlach, & Penke, 2018; Puts 2005; 2006), while vocal attractiveness correlates positively with rated conscientiousness and negatively with rated neuroticism (Zuckerman, Miyake, & Elkin, 1995). These social evaluations and personality judgments based on other peoples' voices are characterized by a high level of agreement between perceivers across different speech contents and contexts (Mahrholz, Belin, & McAleer, 2018; McAleer et al., 2014; Scherer, 1972). Interestingly, three studies suggest that judgments of extraversion and emotional stability based on voice are somewhat accurate (compared with target people's self-reported personality traits; Borkenau & Liebler, 1992, with $N = 100$ self-reported personality; Scherer, 1972, with $N = 59$ self- and peer-reported personality; Scherer, 1978, with $N = 24$ peer-reported personality).

While there are some studies on personality judgments from voices, literature on vocal characteristics and their actual link to target personality and individual differences is rather scarce. Only three early studies have reported direct associations between some vocal characteristics and personality trait variables: Allport and Cantril (1934) reported that more extraverted people had 'louder, more boisterous and carefree voices' (in $N = 3$ male speakers scoring low, medium and high on judge-rated extraversion, respectively). Mallory and Miller (1958) reported that participants, whose voices were judged as "especially high", self-reported lower dominance and higher introversion (in total $N = 372$ without any information on their sex, whereas it remains unclear how many participants had "especially high" voices and were, thus, part of the analyses). Borkenau and Liebler (1992) reported that self-reported agreeableness, but none of the other Big 5 personality traits, correlated significantly with other-rated higher voice pitch (in $N = 100$ with 50% women). Moreover, a relationship between lower voice pitch and markers associated with more self-reported unrestricted sociosexual behavior has been reported in that lower voice pitch in men, as well as more other-rated attractive voices in both sexes, were associated with self-reporting a higher number of sex partners (Hughes, Dispenza, & Gallup, 2004; Puts, 2005).

Nevertheless, no study has directly investigated links between actually measured acoustic parameters and self-reported personal-

ity traits. Therefore, this study's aim is to examine which vocal characteristics, if any, are related to self-reported personality traits. For this purpose, we will combine different independent datasets from previous studies for secondary data analysis, resulting in a large sample size to investigate the relationship between vocal characteristics and personality traits.

1.2. Hypotheses

Based on previous studies, we hypothesize that voice pitch is a valid cue to the speaker's self-reported personality traits. More precisely, as previous studies reported associations between subjectively judged voice pitch and self-reported dominance and extraversion (Mallory & Miller, 1958), agreeableness (Borkenau & Liebler, 1992), and number of sexual partners (Hughes et al., 2004; Puts, 2005), we hypothesize:

Hypothesis 1: Participants with lower voice pitch will self-report as higher on dominance.

Hypothesis 2: Participants with lower voice pitch will self-report as higher on extraversion.

Hypothesis 3: Participants with lower voice pitch will self-report as lower on agreeableness.

Hypothesis 4: Participants with lower voice pitch will self-report as higher (more unrestricted) on sociosexual behavior.

Previous research has mostly neglected potential sex differences in the association between voice pitch and personality traits. As voice pitch is highly sexually dimorphic (sex differences in voice pitch are about 5 SDs; Puts et al., 2012), it seems plausible, that effects might go in opposite directions for male and female speakers. Thus, we decided to analyze speaker's sex as potential moderator variable. We will also add formant frequencies as a predictor variable in an exploratory manner. Formant frequencies have been reported to influence vocal attractiveness and to be another important variable that might influence social impressions (e.g. Collins, 2000; Feinberg et al., 2011), but have not yet been linked to personality. Further, we will perform a number of exploratory analyses investigating the relationships between voice pitch, formant frequencies and conscientiousness, neuroticism and openness for experiences. We decided to do these analyses in an exploratory manner, as there is evidence that judgments of these self-reported traits are somewhat accurate (e.g. Borkenau & Liebler, 1992), and voice parameters are used by others to form these judgments (Aronovitch, 1976; Scherer, 1978; Zuckerman et al., 1995). However, no study so far has investigated whether voice pitch is, indeed, a valid cue to people's conscientiousness, neuroticism or openness for experiences. Additionally, we will also investigate the associations of voice parameters with sociosexuality (full scale score), and its other two facets besides behavior, namely attitude and desire (Penke & Asendorpf, 2008). Again, judgments of sociosexuality at zero acquaintance have been reported to be somewhat accurate (Stillman & Maner, 2009), and the three facets are inter-correlated (Penke & Asendorpf, 2008), but research on voice pitch as a valid cue to sociosexuality and the three facets is missing in the literature.

2. Methods

2.1. Participants

A total of 2217 participants ($n = 918$ men; $n = 1299$ women; aged 18 to 56 years) were recruited in eleven different, independent previous studies focused on other research questions (see Tables 1 and 2 for more information). As the self-reported person-

Table 1

Information about the combined datasets on sample sizes, content of voice recordings, used personality questionnaires and publications in which the sample has been used before and purpose of the original study for which the data has been collected.

Dataset number	N (male/female)	Voice recordings	SOI-R	BFI	NEO-FFI	IAL/IAS-R	Sample characteristics detailed in:	Purpose/topic
1	400 (0/400)	Vowels	X				Jones et al. (2018)	Ovulatory cycle effects and hormonal contraception
2	383 (190/193)	Counting 1–10	X	X	X		Asendorpf, Penke, and Back (2011)	Speed Dating
3	285 (141/144)	Counting 1–10	X	X	X		Penke and Asendorpf (2008 Study 2)	Creating a revised version of the sociosexual orientation inventory
4	265 (0/265)	German rainbow passage	X	X			Stern, Kordsmeyer, and Penke (2021)	Ovulatory cycle effects
5	187 (61/126)	Rainbow passage	X				Puts et al. (2016)	Sexual selection on voice pitch
6	186 (186/0)	German rainbow passage	X				Schild et al. (2020)	The relationship between voice pitch and trustworthiness
7	165 (165/0)	Self- presentation	X	X	X		Kordsmeyer and Penke (2019)	Hormones, personality and mating
8	157 (0/157)	German rainbow passage	X	X	X		Jünger et al. (2018a; 2018b Study 2)	Ovulatory cycle effects
9	133 (43/90)	Standardized sentence		X			N/A	Association of personality and vocal parameters
10	88 (88/0)	Vowels		X			Driebe et al. (2020)	Intelligence and attractiveness
11	65 (65/0)	Rainbow passage	X				Hill et al. (2013)	Sexual selection of men's traits (dominance and attractiveness)

Table 2

Additional information about the combined datasets on age span, dropouts, country in which the data has been collected.

Dataset number	N (male/female)	Age span	Country	Dropouts	Reasons for dropouts
1	400 (0/400)	18–30 years	Scotland	n = 61 dropped out	Did not fill out the SOI-R
2	383 (190/193)	18–54 years	Germany	n = 1 woman	Did not fill out the questionnaires
3	285 (141/144)	19–30 years	Germany	n = 1 woman	Technical problems
4	265 (0/265)	18–35 years	Germany	none	N/A
5	187 (61/126)	18–27 years	USA	none	N/A
6	186 (186/0)	18–56 years	Germany	n = 2 dropped out	Did not fill out the questionnaires
7	165 (165/0)	18–34 years	Germany	n = 1 dropped out	Did not want his recording to be used for further analyses
8	157 (0/157)	18–35 years	Germany	n = 15 dropped out for analyses including the BFI and SOI-R	Did not fill out these questionnaires
9	133 (43/90)	Undergraduates at McMaster University	Canada	n = 24 dropped out	Did not fill out the questionnaire (n = 19), or voice recording missing (n = 5)
10	88 (88/0)	19–31 years	Germany	none	N/A
11	65 (65/0)	18–23 years	USA	n = 9 dropped out	Did not fill out the questionnaire or technical problems

ality traits differ across studies (see Table 1), the sample size for analyses of the different self-reported personality traits are as follows: dominance with $N = 988$ ($n = 492$ women), Big 5 (openness, conscientiousness, extraversion, agreeableness, neuroticism) with $N = 1433$ ($n = 817$ women), and sociosexuality with $N = 2004$ ($n = 1208$ women). The sample sizes for the current analyses greatly exceed sample sizes from previous studies on voices and personality. Thus, we expect a much higher test power for finding effect sizes comparable to previous studies, or even smaller effects.

2.2. Measures

2.2.1. Personality variables

All personality measures were taken via self-report questionnaires: Sociosexuality was measured as a full scale score of the three facets behavior, attitude and desire with the SOI-R (Penke

& Asendorpf, 2008). Example items for the three facets of sociosexuality are “*With how many different partners have you had sex within the past 12 months?*” for sociosexual behavior, “*Sex without love is OK.*” for sociosexual attitude, and “*How often do you have fantasies about having sex with someone with whom you do not have a committed romantic relationship?*” for sociosexual desire (for details see Penke & Asendorpf, 2008). Dominance was measured with the Interpersonal Adjective List (IAL; Jacobs & Scholl, 2005) or the German version of the revised Interpersonal Adjective Scale (IAS-R; Ostendorf, 2001).

Neuroticism, extraversion, openness to experience, agreeableness and conscientiousness were assessed as the dimensions of the Five Factor Model (FFM) of Personality and measured with the German NEO-FFI (Borkenau & Ostendorf, 1993; 12 items per dimension; Sample 2), the BFI-10 (Rammstedt & John, 2007; Sample 9), the BFI 42-item version (Lang, Lüdtke, & Asendorpf, 2001;

Samples 3 and 7), or the BFI 44-item version (John, Donahue, & Kentle, 1991; Samples 4, 8 and 10). Detailed information is shown in Table 1.

2.2.2. Voice recordings and analyses

For voice recordings, participants were instructed to either read an excerpt from a standardized voice passage (e.g. the “rainbow passage”; Fairbanks, 1960; Samples 4, 5, 6, 8 and 11), count from 1 to 10 (Samples 2 and 3), say “A-E-I-O-U” (speaking vowels, Samples 1 and 10), say exactly the same standardized sentence (“Hi, I am a student at McMaster University”, Sample 9), or present themselves (answering the question “What do you think is great about yourself?”, with i.e. “I’m successful at my job”; Sample 7). Detailed information on voice recordings used in the different datasets is shown in Table 1. Length and content of different voice recordings should not affect relationships between personality and vocal parameters, because vocal parameters usually show moderate to strong correlations across different recordings, even independent of length and content (Mahrholz et al., 2018; Puts et al., 2012). Moreover, all recordings are of a neutral content, in which pitch variation is usually very small (Belin, Fillion-Bilodeau, & Gosselin, 2008). To reduce the influence of loudness of different recordings, we normalized loudness to 70 dB for all recordings before analysis. For all samples with multiple voice recordings per participant (due to a within-subjects design with repeated measures), we analyzed recordings from the first session only.

All voices were analyzed using Praat software (Version 6.0.37; Boersma & Weenink, 2018). We measured the following variables: mean F_0 (fundamental frequency), mean formant frequencies (supralaryngeal vocal tract resonances) from which we computed P_f (by standardizing f_1 , f_2 , f_3 and f_4 and dividing the sum by four, following Puts et al., 2012). Scripts for acoustic analyses are publicly available on the OSF (Feinberg, 2018; Puts & Cardenas, 2018).

2.2.3. Data transformations

For F_0 analyses, we used the variable “mean Pitch”, extracted by Praat. P_f was computed using the standardized and aggregated value of the four formants (Puts et al., 2012). All personality measures, F_0 , and P_f were z-standardized prior to analysis.

2.3. Statistical analyses

All analyses were computed with the statistical software R (R Core Team, 2016) and the package brms (Bürkner, 2017) which implements an R interface to the probabilistic programming language Stan (Carpenter et al., 2015). The pre-planned analysis code is publicly available¹ (<https://osf.io/x4jzq>). All data and scripts used for analyses are uploaded to the same OSF page. According to recommendations by Kruschke (2018), we focused on estimating the strength of associations between voice parameters and personality traits. However, to give recommendations for future research, we implemented a decision rule (Makowski & Lüdtke, 2019) where we combined a region of practical equivalence (ROPE) from -0.1 to 0.1 with the 89% highest density interval (HDI) of the estimated effect sizes. This allowed us to differentiate between three scenarios: a) The estimated HDI is completely within the ROPE. Future researchers should not expect to find substantive associations here, unless they think our ROPE was too broad, or can substantially improve on our measurement of voice parameters or dependent variables. b) The HDI overlaps with the ROPE, so we do not know whether the association is outside the ROPE – in other words, our

estimates are insufficiently precise, and future research with larger samples is needed. c) The estimated effect sizes seem substantial because our HDI is outside the ROPE. Future research should work to better characterize these associations.

As all variables except age and sex were standardized, we use weakly informative normal priors (centered on zero, with standard deviations of 3) for all population-level effects and the brms default priors for all other parameters. All models were adjusted for age. Sex was investigated as a moderator variable. Sex was effect-coded without weights, with women coded as -1 and men as 1 . Age was adjusted as a linear effect. For one sample, we lacked precise age information; it was recorded only that participants were undergraduates. Therefore, we used an errors-in-variables approach. In datasets where age was measured, age was entered as measured with a standard error of 0.5 (rounding error). In the undergraduate dataset with unmeasured ages, we assumed an age of 20 with a standard error of 3. In effect, this means this dataset did not contribute much to estimating the age effect, but did not have to be excluded because of missing data. Because the personality traits were measured with scales of varying length, we conducted a robustness check in which we allowed not only the intercepts and slopes to vary by study, but also the residual variation.

We always fitted one “simple” model per personality trait to be predicted. In Wilkinson notation, the model was specified as²:

$$\text{Personality_outcome} = F_0 + P_f + \text{sex} + \text{me}(\text{age}, \text{age_se}) + (1|\text{dataset})$$

To diagnose nonlinearity, we graphed the bivariate relationships between all vocal parameters and all traits in scatterplots overlaid with thin-plate spline smooths (Wood, Pya, & Säfken, 2016). If visual diagnosis indicated nonlinearity or interactions with sex for certain parameter-outcome combinations, we fitted models allowing nonlinearity via thin-plate splines and/or interactions, respectively. If approximative leave-one-out cross-validation (LOO-IC; Vehtari, Gelman, & Gabry, 2016) favored these adapted models over the simple main effect model (LOO-IC lower than by more than 2 standard errors), then we discuss these models instead. We did not apply the HDI + ROPE decision rule to nonlinear effects, but simply showed them visually and discussed them. If we found that the HDI was not within the ROPE or if we found evidence for nonlinear effects, then we also fitted an additional model to see whether this association was invariant across datasets by allowing the relevant linear terms in the regression to vary by dataset and comparing models’ LOO-ICs. In the case of model non-convergence, we first increased the number of iterations and the adapt_delta parameter, as advised by the brms package diagnostics. If this was not sufficient, then we further examined the reasons for non-convergence and potentially set more informative priors or reduce model complexity if absolutely necessary.

3. Results

3.1. Preliminary analyses and deviations from the registered plan

As a first step, we computed Cronbach’s alphas from the available raw data for all personality scales or extracted them from previously published studies using the exact same data. Alphas were predominantly good, thus items were averaged for each scale (Tables 3 and 4). Next, we analyzed the dropout rate for all data-

¹ The analysis code has been produced using simulated data. It only included analyses for the relationship between vocal characteristics and dominance, but analyses for all other personality traits were planned to be identical, but with the respective trait as outcome.

² Age was not assessed in Sample 9. For this sample, we have only the information that all participants were undergraduate students. We opted to include this dataset, sampling from a parametric distribution for age in an errors-in-variables approach for this sample, as reflected by the $\text{me}(\text{age}, \text{age_se})$ term.

Table 3
Cronbach's alphas for sociosexuality and dominance scales.

Dataset number	N (male/female)	SOI-R	SOI-R attitudes	SOI-R behavior	SOI-R desire	IAL/IAS-R dominance	Sample characteristics detailed in:
1	400 (0/400)	$\alpha = 0.85$	$\alpha = 0.81$	$\alpha = 0.82$	$\alpha = 0.85$		Jones et al. (2018)
2	383 (190/193)	M: $\alpha = 0.84$, F: $\alpha = 0.83$	M: $\alpha = 0.81$, F: $\alpha = 0.81$	M: $\alpha = 0.84$, F: $\alpha = 0.81$	M: $\alpha = 0.83$, F: $\alpha = 0.79$	M: $\alpha = 0.82$, F: $\alpha = 0.79$	Asendorpf et al. (2011)
3	285 (141/144)	M: $\alpha = 0.83$, F: $\alpha = 0.83$	M: $\alpha = 0.87$, F: $\alpha = 0.83$	M: $\alpha = 0.85$, F: $\alpha = 0.84$	M: $\alpha = 0.86$, F: $\alpha = 0.85$	$\alpha = 0.75$	Penke and Asendorpf (2008 Study 2)
4	265 (0/265)	$\alpha = 0.85$	$\alpha = 0.83$	$\alpha = 0.84$	$\alpha = 0.78$		Stern et al. (2021)
5	187 (61/126)	$\alpha = 0.84$	$\alpha = 0.84$	$\alpha = 0.81$	$\alpha = 0.84$		Puts et al. (2016)
6	186 (186/0)	$\alpha = 0.85$	$\alpha = 0.81$	$\alpha = 0.85$	$\alpha = 0.84$		Schild et al. (2020)
7	165 (165/0)	$\alpha = 0.81$	$\alpha = 0.71$	$\alpha = 0.86$	$\alpha = 0.85$	$\alpha = 0.74$	Kordsmeyer and Penke (2019)
8	157 (0/157)	$\alpha = 0.85$	$\alpha = 0.88$	$\alpha = 0.86$	$\alpha = 0.76$	$\alpha = 0.78$	Jünger et al. (2018a; 2018b Study 2)
9	133 (43/90)						N/A
10	88 (88/0)						Driebe et al. (2020)
11	65 (65/0)	$\alpha = 0.81$	$\alpha = 0.81$	$\alpha = 0.79$	$\alpha = 0.79$		Hill et al. (2013)

Table 4
Cronbachs alphas for Big 5 personality trait scales.

Dataset number	N (male/female)	Neuroticism	Extraversion	Openness to experience	Agreeableness	Conscientiousness	Sample characteristics detailed in:
1	400 (0/400)						Jones et al. (2018)
2	383 (190/193)	M: $\alpha = 0.86$, F: $\alpha = 0.83$	M: $\alpha = 0.79$, F: $\alpha = 0.74$	M: $\alpha = 0.71$, F: $\alpha = 0.65$	M: $\alpha = 0.75$, F: $\alpha = 0.74$	M: $\alpha = 0.83$, F: $\alpha = 0.81$	Asendorpf et al. (2011)
3	285 (141/144)	$\alpha = 0.79$	$\alpha = 0.84$	$\alpha = 0.81$	$\alpha = 0.65$	$\alpha = 0.78$	Penke and Asendorpf (2008 Study 2)
4	265 (0/265)	$\alpha = 0.82$	$\alpha = 0.85$	$\alpha = 0.83$	$\alpha = 0.75$	$\alpha = 0.82$	Stern et al. (2021)
5	187 (61/126)						Puts et al. (2016)
6	186 (186/0)						Schild et al. (2020)
7	165 (165/0)	$\alpha = 0.81$	$\alpha = 0.87$	$\alpha = 0.78$	$\alpha = 0.73$	$\alpha = 0.84$	Kordsmeyer and Penke (2019)
8	157 (0/157)	$\alpha = 0.77$	$\alpha = 0.84$	$\alpha = 0.82$	$\alpha = 0.74$	$\alpha = 0.80$	Jünger et al. (2018a; 2018b Study 2)
9	133 (43/90)	NA	NA	NA	NA	NA	N/A
10	88 (88/0)	$\alpha = 0.83$	$\alpha = 0.90$	$\alpha = 0.81$	$\alpha = 0.76$	$\alpha = 0.85$	Driebe et al. (2020)
11	65 (65/0)						Hill et al. (2013)

Note: Raw data for dataset 9 was not available to compute Cronbachs alphas.

sets. Dropout rates differed from the expected rates for some datasets, leading to a loss of 2.78% of all participants. As a consequence, we had to update Table 2 and the total number of participants in the methods part. We note that this leads to a slight deviation from our registered plan³. We then aggregated data across datasets. Voice parameters were not transformed prior to aggregation, but we computed the percentage of the maximum possible to aggregate personality variables with different response scales. We found that voice recordings of one dataset (Dataset 10) had a low quality, but filtering background noise (Jadoul, Thompson, & de Boer, 2018) led to implausible values for formant frequencies and the pattern of results was virtually identical when dataset 10 was omitted. Hence, we report results for unfiltered voice recordings to keep voice analysis procedures identical. Further, we corrected two mistakes in our registered statistical analysis part, because of discrepancies between the written text and the registered R script. More precisely, first, we wrote that we will use 95% HDIs in the text but modelled 89% HDIs

³ Sample sizes as reported in the Stage 1 Registered Report: Total of 2278 participants ($n = 931$ men; $n = 1347$ women; aged 16 to 56 years) was changed to 2217 participants ($n = 918$ men, $n = 1299$ women; aged 18 to 56 years) in the Stage 2 Registered Report. The sample sizes for analyses of the different self-reported personality traits were changed as follows: Big 5 (openness, conscientiousness, extraversion, agreeableness, neuroticism) from $N = 1449$ ($n = 822$ women) to $N = 1433$ ($n = 817$ women), and sociosexuality from $N = 2082$ ($n = 1283$ women) to $N = 2004$ ($n = 1208$ women). The sample size for dominance analyses remained identical.

in the script. We decided to report 89% HDIs, as recent guidelines recommend using them rather than 95% HDIs, because the latter might lack stability (Kruschke, 2014; McElreath, 2018). The value 89% rather than 90% has been recommended by McElreath to emphasize the arbitrariness of specific credibility limits. For our estimates, we also report 95% HDIs, so readers can easily assess these against our ROPE as well. Second, we updated the Wilkinson notation formula, as it did not reflect the described simple model, but rather a more complex model including terms for an interaction between voice parameters and sex. However, we did visually inspect the data for interactions with sex, as described in the text, and where interactions seemed possible, we conducted a model comparison and equivalence tests. In addition, we decided to still investigate potential interaction effects for our preregistered outcomes and describe the results at the beginning of the exploratory analyses part. Details are described below or can be found at the supplementary website (https://rubenarlan.github.io/voices_personality/).

Next, we computed intercorrelations between all variables. Results for all preregistered variables are displayed in Table 5, intercorrelations for all variables can be found at the supplementary website (or in the codebook on the OSF page).

Then, as preregistered, we visually inspected bivariate relationships between vocal parameters and all traits to diagnose nonlinearity or interactions with sex. Visual diagnoses support nonlinear effects and interactions with sex for several outcomes,

Table 5
Correlations between sex, age, voice pitch, and preregistered outcomes.

	Sex	Age	Voice pitch	Dominance	Extraversion	Agreeableness	SOI-R behavior
Sex	1.00	0.00	0.00	0.00	0.00	0.00	0.00
Age	0.20	1.00	-0.12	-0.12	-0.03	0.11	0.30
Voice pitch	-0.93	-0.23	1.00	-0.04	-0.06	-0.02	-0.11
Dominance	-0.03	-0.13	0.01	1.00	0.51	-0.23	0.13
Extraversion	-0.09	-0.04	0.06	0.51	1.00	0.11	0.20
Agreeableness	-0.13	0.09	0.11	-0.22	0.12	1.00	0.01
SOI-R behavior	-0.01	0.29	-0.03	0.13	0.21	0.01	1.00

Note: Correlations above the diagonal are after partialling out effects of sex.

but multivariate, age-adjusted models including these terms usually did not fit better by our criterion of 2 standard errors of the LOO-IC. For agreeableness, we found that voice pitch had a positive effect for men, but not for women, so that men with deeper voices were less agreeable. A model including this interaction, however, fell short of our LOO-IC criterion. We found no compelling evidence for nonlinear associations with voice parameters after adjusting for age and sex, but age had a nonlinear association that differed by sex with sociosexual behavior. For testing our preregistered hypotheses, as well as for exploratory analyses, we fitted multiple models and compared their fit. These models included: a) a simple model with the self-reported personality trait as an outcome, voice pitch, formant position, sex and age as fixed effect predictors and a random intercept for dataset b) a model that added terms after visual inspection (e.g. nonlinear effects, interactions with sex) c) a model, in which the effect for voice pitch and sex were allowed to vary across datasets. More complex models (e.g. including interaction effects, nonlinear effects, or effects varying by dataset) never led to a better model fit by our 2 LOO-IC SE criterion. Hence, we report the simple model, as planned. However, results for all models and comparisons can be found at our supplementary website (and on the OSF page).

3.2. Preregistered hypotheses

3.2.1. Dominance

First, we tested whether participants with lower voice pitch self-report as higher on dominance (Hypothesis 1). The data are consistent with a substantial negative linear relationship between voice pitch and dominance, (estimate = -0.27, 95% CI [-0.45; -0.08]) and support Hypothesis 1 (see Fig. 1). Moreover, men reported substantially lower dominance than women (estimate = -0.30, 95% CI [-0.52; -0.08]). The 89% HDI for voice pitch and for sex fall entirely outside the ROPE (see Fig. 2), but the 89% HDI for formant position and age falls almost entirely within it. This evidence is consistent with a non-negligible association, where people with deeper voices are more dominant (after adjusting for age and sex), and women self-reported to be more dominant than men. Further research is needed to verify whether the association with formant position (estimate = -0.00, 95% CI [-0.13; 0.12]) is truly negligible in size.

3.2.2. Extraversion

Second, we tested whether participants with lower voice pitch self-report as higher on extraversion (Hypothesis 2). The data are consistent with a substantial negative linear relationship between voice pitch and extraversion (estimate = -0.23, 95% CI [-0.38; -0.08]). Thus, we find support for Hypothesis 2 (see Fig. 1). Moreover, men reported substantially lower extraversion (estimate = -0.28, 95% CI [-0.46; -0.10]). The 89% HDI for voice pitch and for sex fall entirely outside the ROPE (see Fig. 2), but the 89% HDI for formant position and age fall about half within it. This evidence is consistent with a non-negligible association, where people with deeper voices are more extraverted (after

adjusting for age and sex), and women self-reported to be more extraverted than men. Further research is needed to verify whether the association with formant position (estimate = 0.08, 95% CI [-0.03; 0.18]) is truly negligible in size.

3.2.3. Agreeableness

Third, we tested whether participants with lower voice pitch self-report as lower on agreeableness (Hypothesis 3). The credible intervals for voice pitch (estimate = 0.03, 95% CI [-0.12; 0.18]), formant position (estimate = 0.03, 95% CI [-0.07; 0.13]), and sex (estimate = -0.00, 95% CI [-0.18; 0.18]) include zero, but the 89% HDI does not fall entirely within the ROPE (see Fig. 2). This evidence is consistent with a negligible association between voice pitch (see Fig. 1), formant position, sex and agreeableness and does not support Hypothesis 3. Further research is needed to verify whether the associations are truly negligible in size. As reported above, when visually inspecting the data for agreeableness, we detected an interaction effect between voice pitch and sex, in that the relationship was negative for women (women with deeper voices were more agreeable), but rather positive for men (men with deeper voices were less agreeable). However, the 89% HDI for this interaction effect overlaps the ROPE (2% in the ROPE), and the model fit was not superior by our criterion of two LOO-IC standard errors to the simple model reported above. Thus, future research is needed to verify whether the interaction effect is negligible in size or not.

3.2.4. Sociosexual behavior

Fourth, we tested whether participants with lower voice pitch self-report as higher (more unrestricted) on sociosexual behavior (Hypothesis 4). The data are consistent with a substantial negative linear relationship between voice pitch and sociosexual behavior (estimate = -0.21, 95% CI [-0.33; -0.09]) and support Hypothesis 4 (see Fig. 1). Moreover, men reported substantially more restricted sociosexual behavior than women did (estimate = -0.24, 95% CI [-0.38; -0.09]). The 89% HDI for voice pitch, sex and age fall entirely outside the ROPE (see Fig. 2), whereas the 89% HDI for formant position falls entirely within it. This evidence is consistent with a non-negligible association, where people with deeper voices report more unrestricted sociosexual behavior (after adjusting for age and sex), and women self-reported more unrestricted sociosexual behavior than men did. However, according to our criterion the association with formant position is negligible (estimate = -0.02, 95% CI [-0.10; 0.07]).

3.3. Exploratory analyses and robustness checks

Although our model comparisons using approximative cross-validation never led us to conclude the presence of an interaction between voice pitch and sex by the criterion we set, we decided to also test the HDIs for all interaction effects for our preregistered effects against our ROPE to be able to give recommendations for future research. Further, we test relationships between all other self-reported personality traits (neuroticism, openness, conscien-

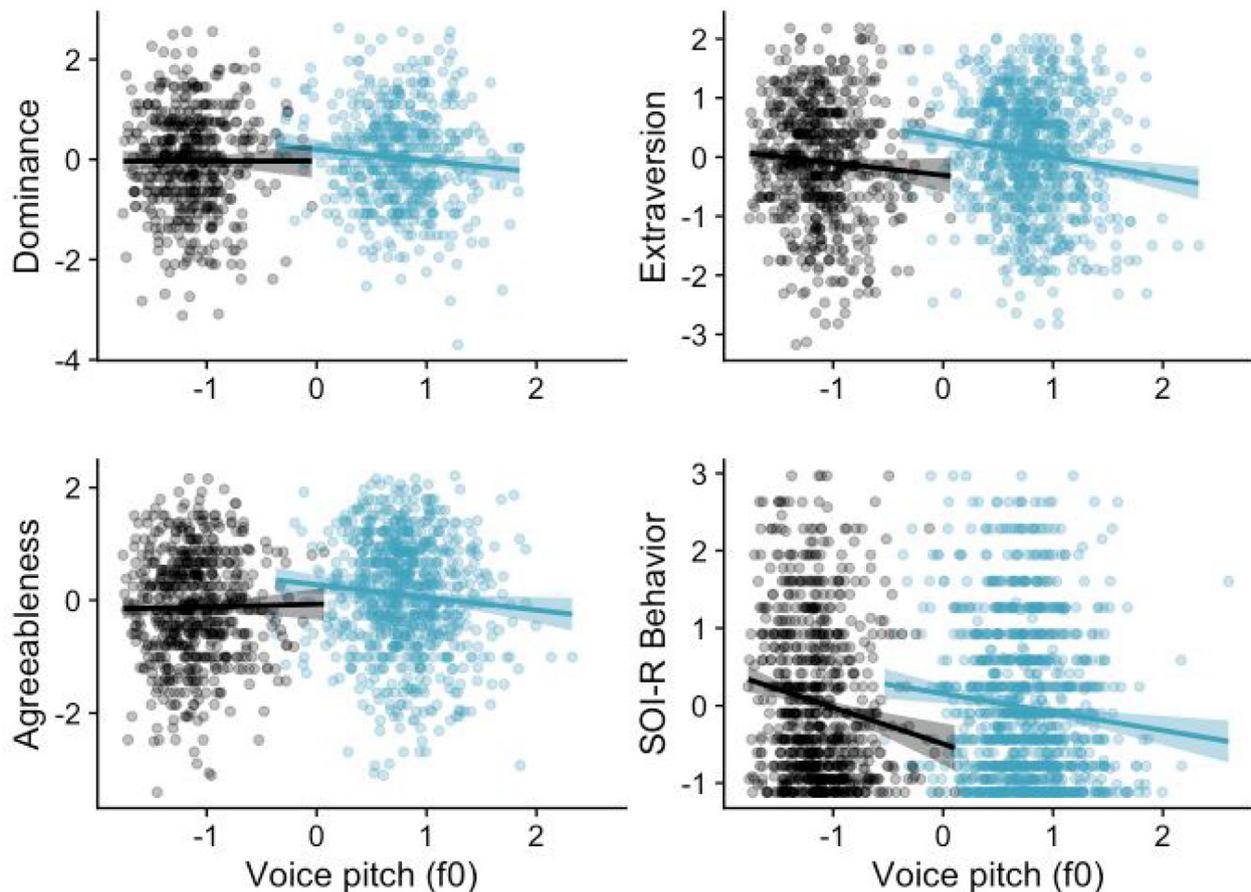


Fig. 1. Results for Bayesian multilevel models including dominance, extraversion, agreeableness and sociosexual behavior as a function of voice pitch and male (black) or female (blue) speakers. *Note:* All personality variables and voice pitch were z-scored. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

tiousness, global sociosexual orientation, sociosexual attitudes, and sociosexual desire) and voice pitch or formant position, adjusting for sex and age, equivalent to the preregistered hypotheses. Last, as extraversion and dominance were correlated (Table 5), we adjust the above reported dominance analyses for extraversion and vice versa.

3.3.1. Interaction effects of voice pitch and sex

Although visual diagnoses and model fit criteria suggested that voice pitch and sex do not seem to show substantial interaction effects for our preregistered outcomes, we also tested for potential interaction effects and did separate analyses for men and women. The reasons for these additional analyses are that a) voice pitch is highly sexually dimorphic and b) we stated we would investigate interaction effects in our registered plan (although we followed our plan to report the simple model based on visual diagnoses and model fit). Briefly, all equivalence tests for an interaction between voice pitch and sex for our four preregistered outcomes were undecided, that is 89% HDIs always overlapped our ROPE. The HDI excluded zero only for agreeableness. For agreeableness, the interaction pattern of the association in men was opposite to that in women: men with deeper voices were less agreeable, whereas women with deeper voices were more agreeable. However, only the 95% HDI for the interaction excluded zero, the 95% HDIs for main effects of voice pitch estimated separately by sex both included zero and 2% of the 89% HDI of the interaction fell within the ROPE. For dominance, and extraversion interaction effects were positive, reflecting a weaker negative association

between voice pitch and the outcome in men than in women, but 77% and 88% of the interaction HDIs fell inside the ROPE, respectively. For sociosexual behavior, the interaction effect was negative, so that the negative association between voice pitch and sociosexual behavior was slightly stronger in men than in women, but again 83% of the HDI fell inside the ROPE. Details can be found at our supplementary website. Future research should investigate whether these interactions are negligible.

3.3.2. Neuroticism

For neuroticism, the data are consistent with a positive linear relationship between voice pitch and neuroticism (estimate = 0.16, 95% CI [0.02; 0.31]). The 89% HDI for voice pitch excludes zero but is partly inside the ROPE (see Fig. 3). 86% of the HDI for formant position (estimate = -0.05, 95% CI [-0.15; 0.05]) and 27% for sex (estimate = -0.15, 95% CI [-0.32; 0.03]) fall within the ROPE. This evidence is consistent with an association where people with deeper voices have lower neuroticism (after adjusting for age and sex), but further research is needed to test whether the associations with voice pitch, formant position and sex are negligible in size.

3.3.3. Openness to experience

For openness to experience, the data are consistent with a negative linear relationship between sex and openness to experience (estimate = -0.22, 95% CI [-0.42; -0.04]). This evidence is consistent with an association, in that women self-report higher openness to experience. Further, the data are consistent with a negative linear relationship between voice pitch and openness to experience

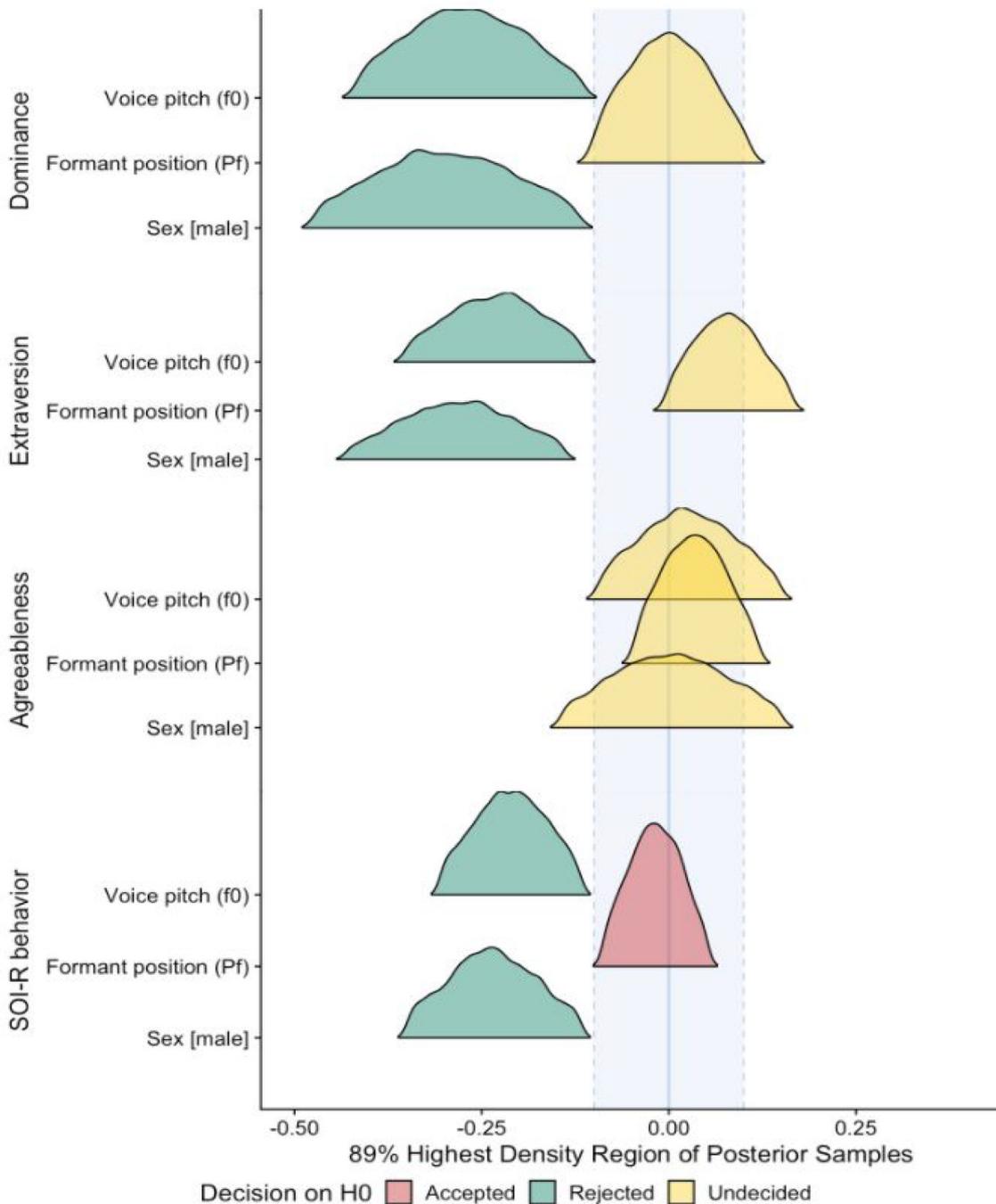


Fig. 2. Results for the preregistered hypotheses: 89% highest density intervals for dominance, extraversion, agreeableness and sociosexual behavior as a function of voice pitch, formant position and male sex. *Note:* The figure shows the 89% highest density intervals (HDI) for the estimated linear effects of voice pitch (F_0), formant position (P_f), and male sex across all (preregistered) outcomes. The shaded blue region shows the region of practical equivalence (ROPE). For dominance, extraversion and sociosexual behavior, we find that the 89% HDI lies completely outside the ROPE and reject the null hypothesis of no association between the outcomes and voice pitch. For formants, we never reject the null hypothesis, but because the HDI always overlaps the ROPE boundaries, we cannot conclude that voice pitch has only negligible associations either. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(estimate = -0.17 , 95% CI [-0.32 ; -0.01]). The HDIs for sex and voice pitch exclude zero but are partly (sex: 5%, voice pitch: 14%) inside the ROPE (see Fig. 3). The HDI for formant position (estimate = 0.00 , 95% CI [-0.10 ; 0.11]) falls entirely within the ROPE. This evidence is consistent with an association where people with deeper voices have higher openness to experience (after adjusting for age and sex), but further research is needed to test whether the associations with voice pitch and sex are negligible in size.

3.3.4. Conscientiousness

For conscientiousness, the 89% HDI for voice pitch (estimate = -0.02 , 95% CI [-0.17 ; 0.13]) and formant position (estimate = 0.04 , 95% CI [-0.06 ; 0.14]) fall almost entirely inside the ROPE (see Fig. 3). The 89% HDI for sex is 53% inside the rope (estimate = -0.09 , 95% CI [-0.27 ; 0.09]). Further research is needed to test whether the associations with voice pitch, formant position and sex are negligible in size.

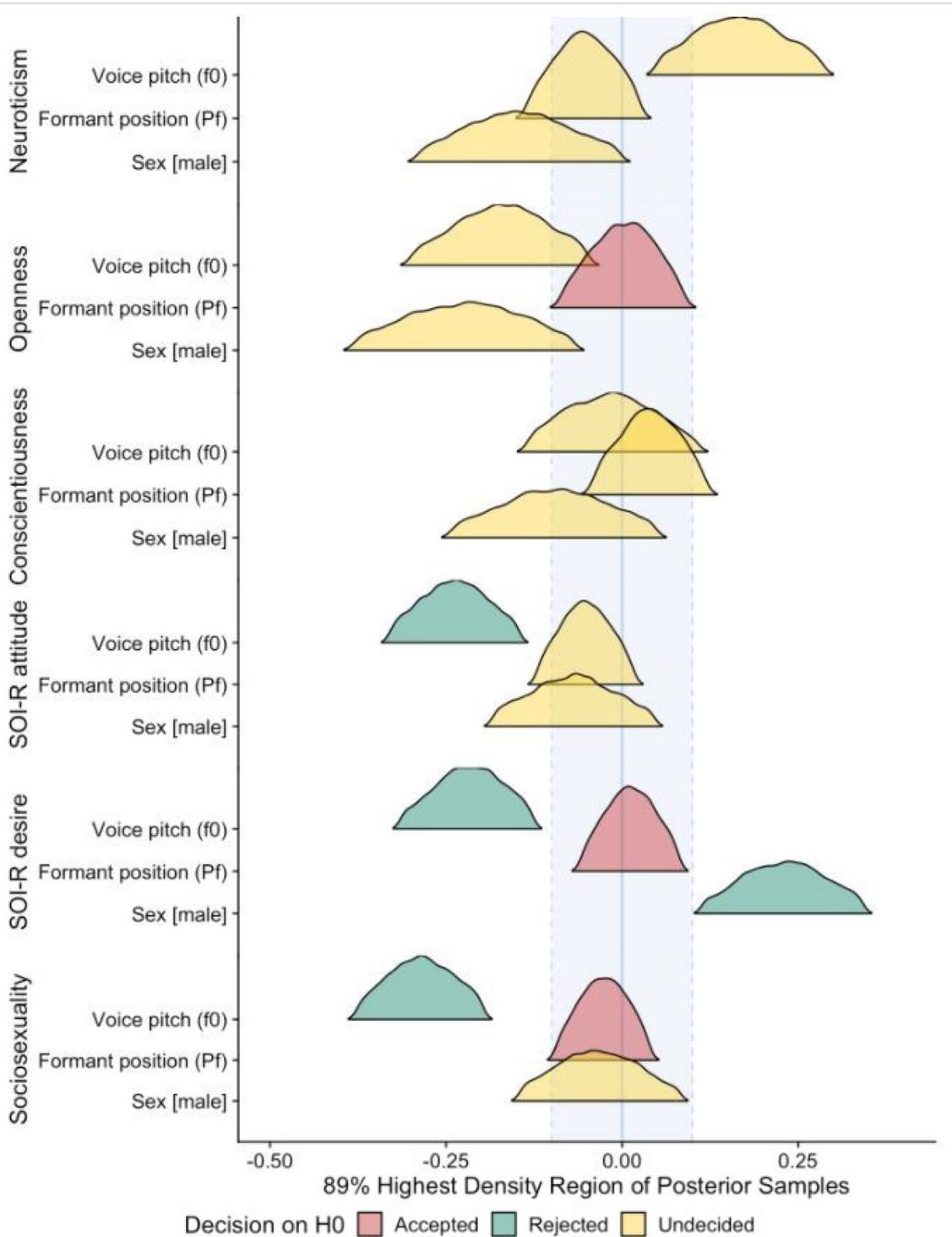


Fig. 3. Results for the exploratory analyses: 89% highest density intervals for neuroticism, openness to experience, conscientiousness, sociosexual attitude, desire and orientation, as a function of voice pitch, formant position and male sex. *Note:* The figure shows the 89% highest density intervals (HDI) for the estimated linear effects of voice pitch (F_0), formant position (P_f), and male sex across all (exploratory) outcomes. The shaded blue region shows the region of practical equivalence (ROPE). For sociosexual attitude, desire and orientation, we find that the 89% HDI lies completely outside the ROPE and reject the null hypothesis of no association between the outcomes and voice pitch. For formants, we never reject the null hypothesis, but because the HDI always overlaps the ROPE boundaries, we cannot conclude that voice pitch has only negligible associations either. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.3.5. Sociosexual orientation

For sociosexual orientation, the data are consistent with a substantial negative linear relationship between voice pitch and unrestricted sociosexuality. The 89% HDI for voice pitch falls entirely outside the ROPE (estimate = -0.28 , 95% CI [-0.39 ; -0.17]). The

89% HDI for formant position (estimate = -0.03 , 95% CI [-0.11 ; 0.06]) falls entirely within the rope. For sex, (estimate = -0.04 , 95% CI [-0.18 ; 0.10]) the 89% CI falls almost entirely within it (see Fig. 3). This evidence is consistent with a non-negligible association, where people with deeper voices have a less restricted

sociosexuality (after adjusting for age and sex), and a negligible association with formant position.

3.3.6. Sociosexual attitude

For sociosexual attitude, the data are consistent with a substantial negative linear relationship between voice pitch and unrestricted sociosexual attitude. The 89% HDI for voice pitch falls entirely outside the ROPE (estimate = -0.24 , 95% CI [-0.35 ; -0.12]), but the 89% HDI for formant position (estimate = -0.05 , 95% CI [-0.14 ; 0.03]) and sex (estimate = -0.07 , 95% CI [-0.21 ; 0.07]) fall almost entirely within it (see Fig. 3). This evidence is consistent with a non-negligible association, where people with deeper voices have less restricted sociosexual attitudes (after adjusting for age and sex). Further research is needed to test whether the associations with formant position and sex are negligible in size.

3.3.7. Sociosexual desire

For sociosexual desire, the data are consistent with a substantial negative linear relationship between voice pitch and unrestricted sociosexual desire, as well as a substantial positive relationship between sex and sociosexual desire. The 89% HDIs for voice pitch (estimate = -0.22 , 95% CI [-0.33 ; -0.10]) and for sex (estimate = 0.23 , 95% CI [0.09 ; 0.37]) fall entirely outside the ROPE (see Fig. 3). The 89% HDI for formant position (estimate = 0.01 , 95% CI [-0.07 ; 0.10]) falls entirely within it. This evidence is consistent with a non-negligible association, where people with deeper voices have less restricted sociosexual desire (after adjusting for age and sex), and men self-report higher sociosexual desire than women, and a negligible association with formant position.

3.3.8. Extraversion adjusting for dominance and vice versa

As extraversion and dominance were correlated ($r = 0.51$), we decided to repeat the analysis for extraversion reported above, adding dominance as a control variable. In this model, the negative linear association between voice pitch and extraversion (estimate = -0.07 , 95% CI [-0.23 ; 0.08]), as well as the association between extraversion and sex (estimate = -0.08 , 95% CI [-0.26 ; 0.11]), decreased in size. When adjusting the dominance model for extraversion (by adding extraversion as a control variable), the negative linear associations between voice pitch and dominance (estimate = -0.14 , 95% CI [-0.30 ; 0.02]) and between dominance and sex (estimate = -0.17 , 95% CI [-0.35 ; 0.02]) decreased less than the above reported relationships of extraversion with voice pitch or sex after adjusting for dominance.

3.4. Robustness checks

As described in our registered plan, we conducted robustness checks for the preregistered analyses in which we fitted models that allowed not only the intercepts and slopes but also the residual variation to vary by study because personality traits were measured with scales of different lengths. For extraversion, agreeableness and sociosexuality as outcomes, these models showed a better model fit than the simple model reported above. However, the pattern of results was virtually identical, except that effect sizes were slightly smaller, and the uncertainty was elevated.

In addition to the planned checks, we checked robustness to alternative arbitrary decisions that arose *post-hoc* during data analysis. Namely, because the voice recordings differed in content and quality, we formed three different datasets in addition to the main dataset: 1. We standardized voice parameters within sample prior to aggregation. In samples of exclusively men or women, we re-applied the average estimated sex differences in means and variances. 2. We removed outliers that were more than 2.5 median absolute deviations from the median for fundamental frequency

or formant position. 3. We omitted sample 1 and 10, in which participants had only spoken vowels. We examined effect sizes for our preregistered tests for each of these alternative datasets. All effect sizes were virtually identical across approaches. Further, we repeated the preregistered analyses substituting individual estimates of formant 1 to 4 for the formant position aggregate. The 95% HDIs for the formants always overlapped zero, with one exception, the effect of formant 3 on agreeableness (estimate = 0.14 , 95% CI [0.01 ; 0.27]). Future research should consider examining individual formants, but otherwise these explanatory analyses should be interpreted with caution. Further details are reported in our online supplement.

4. Discussion

In the current study, we investigated the association between voice pitch and formant position with different self-reported personality traits, adjusted for the effects of sex and age. Results indicate that participants with lower voice pitch self-report as higher on dominance, extraversion and more unrestricted on sociosexual behavior, in line with our Hypotheses 1, 2 and 4. However, Hypotheses 3, which states that participants with lower voice pitch self-report as lower on agreeableness, was not supported. Exploratory analyses suggest that participants with lower voice pitch self-report as lower on neuroticism, higher on openness to experience, and as more unrestricted on sociosexual orientation, sociosexual attitudes and sociosexual desire. However, effects for associations between voice pitch and neuroticism, openness to experience and conscientiousness remain undecided, but the HDI for conscientiousness falls almost entirely inside the ROPE.

We did not find any compelling association between formant position and self-reported personality traits, but future research is needed to verify whether most associations are truly negligible in size. Further, although voice pitch is highly sexually dimorphic (the sex difference in voice pitch is approximately 5 SDs; Puts et al., 2012), we did not find any compelling sex differences in associations between voice pitch (or formant position) and self-reported personality traits. However, we found main effects of sex, in that men self-reported lower dominance, extraversion, and less restricted sociosexual desire, but more restricted sociosexual behavior, as compared to women. Interestingly, many of these associations with sex ran counter to the respective main effect of voice pitch. For instance, even though people with lower voice pitch reported having more partners (sociosexual behavior), and men have much lower voice pitch, men reported fewer partners. The exception to this pattern was sociosexual desire, where men reported higher desire, and people with lower voice pitch also reported higher desire. In effect, this can be interpreted as indicating that voice pitch is a valid cue of personality, conditional on inferring sex. Hence, a man with a high voice pitch can be evaluated differently from a woman with a very similar voice pitch, which would be low for a woman. Because the distributions almost do not overlap, it is possible to characterize voices close together in absolute pitch as “high for a man” or “low for a woman”. In a scatterplot, this shows as a large “step” in the linear regression (Fig. 1). The woman with the lowest voice pitch would be predicted to have less restricted sociosexual behavior than a man with a similar low voice pitch. In the following, we interpret these findings and highlight implications for future research.

4.1. Voice pitch and self-reported Big 5 personality traits

While multiple previous studies reported that a person's voice might be an important cue when forming social impressions, e.g. about other people's personality traits (e.g. Borkenau & Liebler,

1992), evidence was still unclear whether voice pitch might, indeed, be a valid cue for such judgments. Only a small number of previous studies investigated and reported an association between voice pitch and personality traits (Allport & Cantril, 1934; Borkenau & Liebler, 1992; Mallory & Miller, 1958); however, in these studies, voice pitch was subjectively rated, not objectively measured, which engenders method variance and might be confounded with impressions about people's personality. Nevertheless, our results are partly in line with these previous studies in that people with lower voice pitch self-reported higher extraversion (as also reported in Allport & Cantril, 1934; Mallory & Miller, 1958). In contrast, Borkenau and Liebler (1992) reported that people with lower voice pitch self-report lower on agreeableness but found no significant effects for any of the other Big 5 personality traits. Differences in results (we report no compelling evidence for an association of voice pitch and agreeableness, but negative linear effects for extraversion and openness to experience) might, at least in part, be a function of different sample sizes and thus test power. Our reported effects were small in magnitude, and given that our sample size for these analyses ($n = 1433$) greatly exceeded the sample reported by Borkenau and Liebler (1992; $N = 100$), it is plausible that their analyses were underpowered to detect small effects. While we conclude uncertainty for an association between voice pitch and agreeableness, our study was adequately powered to detect the effect size ($r = 0.22$) reported by Borkenau and Liebler (1992), and our 89% HDI for voice pitch excludes the value 0.22. Differences in results might also stem from methodological differences, especially as our study is the first study using objectively analyzed voice recordings, as noted above. The overall evidence of previous work and our study regarding voice pitch and neuroticism, openness to experience or conscientiousness rather seems to suggest no compelling evidence for a substantial association. However, based on our analyses, we conclude uncertainty for these effects and suggest that future research is needed to test whether associations are negligible in size.

4.2. Voice pitch and self-reported dominance

Our results suggest that people with lower voice pitch self-report higher dominance. This finding is in line with previous research reporting a comparable association (Mallory & Miller, 1958). Further, it is in line with the hypothesis that voice pitch honestly signals dominance, power and status (Aung & Puts, 2020). This means that lower voice pitch also imposes costs (e.g., facing retribution when falsely signaling dominance while speaking to competitors that are more formidable), as people with lower voice pitch are also perceived as being more dominant by listeners (Aung & Puts, 2020; Borkowska & Pawlowski, 2011; Collins, 2000; Hodges-Simeon et al., 2010; Puts et al., 2006, 2007). Moreover, it has been reported that, when talking to more dominant or prestigious competitors, men raise their voice (Aung & Puts, 2020; Leongómez, Mileva, Little, & Roberts, 2017; Puts et al., 2006). While we did not investigate dominance perceptions based on voice pitch and their accuracy in our study, our results suggest that perceived associations of dominance and voice pitch, as reported in previous studies, might be accurate.

A potential biological mechanism influencing the relationship between voice pitch and self-reported dominance might be testosterone. More precisely, a recent meta-analysis reported that, in men, lower voice pitch is linked to higher levels of testosterone (Aung & Puts, 2020), whereas the association for women seems to be less clear. Further, one study reported a positive, small-to-medium correlation between testosterone levels and self-reported dominance in men and women (Sellers, Mehl, & Josephs, 2007). However, the latter effect has been reported to be much smaller and potentially not replicable in other studies

(Archer, Graham-Kevan, & Davies, 2005; Kordsmeyer & Penke, 2019). Another mechanism that has previously been discussed might be that lower voice pitch and higher dominance signal better phenotypic condition (e.g. good nutritional status or low pathogen load). Especially men in better phenotypic condition should have longer and denser vocal folds, as well as less anxiety during social interactions (Puts & Aung, 2019). This is potentially due to the effects of higher testosterone and lower cortisol levels on vocal fold anatomy, as well as the influence of autonomic arousal on vocal fold tension. The claim that voice pitch might signal phenotypic condition has been challenged by recent non-replications (reviewed in Feinberg, Jones, & Armstrong, 2019), although meta-analyses indicate that voice pitch is modestly related to measures of condition such as strength, and endocrine profiles (and weakly with body size) in the predicted directions (Aung & Puts, 2020).

Importantly, dominance is a facet of extraversion (McCrae & Costa, 1989; Ostendorf & Angleitner, 2003) and indeed both variables correlated in our datasets ($r = 0.51$). The association between extraversion and voice pitch was descriptively smaller than the association between dominance and voice pitch and reduced more when adjusted for dominance than vice versa, but since dominance was measured with a separate measure and we did not have facet level data for extraversion, which would allow us to estimate associations between voice pitch and non-dominance facets of extraversion, we caution against a strong interpretation of these findings.

4.3. Voice pitch and self-reported sociosexual orientation

Besides reported associations between voice pitch and extraversion, openness to experience and dominance, we report substantial negative relationships between voice pitch and sociosexual orientation, behavior, attitude and desire that are in line with previous research. Previous studies have shown that men with lower voice pitch report higher mating success and a higher number of sexual partners (e.g. Puts, 2005), which are captured in our study by a less restricted sociosexual behavior. These findings are in line with an association of lower voice pitch with a higher probability of self-reported infidelity in committed relationships for men and women, as reported by Schild, Stern, and Zettler (2020, *in press*) in some of the samples that also went into the current analyses. A potential explanation for an association between sociosexual behavior and voice pitch might be attractiveness, in that higher attractiveness potentially increases the opportunity for having a higher number of sexual partners. More precisely, lower voice pitch is perceived as being more attractive (e.g. Collins, 2000; Feinberg et al., 2011; Hodges-Simeon et al., 2010; Jünger, Motta-Mena, & et al., 2018; Puts 2005; 2006). However, this is only evident in men, the association might rather be in the opposite direction for women (Feinberg, DeBruine, Jones, & Perrett, 2008) or there may be little or no relationship between voice pitch and vocal attractiveness in women (Puts et al., 2016). Thus, higher voice attractiveness cannot explain our findings for women. Further, vocal attractiveness does not seem to mediate the association between voice pitch and infidelity in committed relationships (Schild et al., *in press*) and the relationship between voice pitch and physical attractiveness seems to be rather weak (Fisher, Hahn, DeBruine, & Jones, 2016).

An alternative explanation for the association between voice pitch and sociosexuality might be that voice pitch influences sociosexual behavior via its effects on dominance. Indeed, voice pitch seems to have stronger influences on perceptions of dominance than on perceptions of attractiveness (e.g. Borkowska & Pawlowski, 2011; Kordsmeyer, Hunt, Puts, Ostner, & Penke, 2018). Further, both sexes sometimes compete aggressively to obtain and retain mates (Archer, 2009; Campbell, 2004), and suc-

cess in same-sex dominance competition has been found to predict men's number of sexual partners more strongly than their attractiveness to women (Hill et al., 2013; Kordsmeyer et al., 2018). In Bolivian forager-horticulturists, experimentally lowered voice pitch increased perceptions of men's fighting ability among men, but decreased men's attractiveness to women, and men with lower voice pitch tended to have more fertile wives (Rosenfield, Sorokowska, Sorokowski, & Puts, 2020). Thus, signaling dominance via lower voice pitch might indirectly lead to a higher number of sexual partners.

A third explanation might be androgenic masculinization, in that testosterone influences both voice pitch and sociosexuality. Lower voice pitch is an indicator of phenotypic masculinization (e.g. Puts et al., 2012), and it has been reported that a more unrestricted sociosexual orientation and a higher number of sexual partners are linked to higher masculinity in men and women (Burri, Spector, & Rahman, 2015; Ostovich & Sabini, 2004). Testosterone levels might potentially be a common cause of both vocal and sociosexual masculinization (Aung & Puts, 2020). More precisely, lower voice pitch is linked to higher testosterone levels in men (Aung & Puts, 2020) and testosterone administration lowers voice pitch in biological females (Huang et al., 2015; Irwig, Childs, & Hancock, 2017). Further, higher testosterone has been linked to more unrestricted sociosexual orientation in women (Shirazi et al., 2019) and men (but only when cortisol levels are low; Stern et al., 2020; Puts et al., 2015). Yet, it remains unclear whether individual differences in testosterone levels could potentially be a common cause explaining an association between voice pitch and sociosexual orientation (including all three facets). This research question is an interesting direction for future research.

4.4. Formant position

In all of our analyses, the HDI for formant position fall almost or entirely into the predefined ROPE, suggesting no compelling evidence for a link between formant position and any of the investigated self-reported personality traits. One possible explanation for the generally stronger relationships between voice pitch and personality as compared to those between voice pitch and formant position is that average formant frequencies are strongly tied to the length of the supralaryngeal vocal tract, which is relatively fixed in adulthood. By contrast, voice pitch depends on (1) the thickness and density of the vocal folds, which are skeletal muscle that may be more responsive to fluctuations in steroid hormone levels, as well as (2) tension on the vocal folds, which depends on autonomic arousal and hence may be influenced by personality-related emotional responses to social situations. However, we cannot conclude that formant position has negligible effects either, as some values outside the ROPE remain plausible. To our knowledge, no previous study has investigated associations between formant frequencies and self-reported personality traits, although it has been reported that formant frequencies influence vocal attractiveness and thus might influence social impressions (e.g. Collins, 2000; Feinberg et al., 2011). As some of our results suggest uncertainty, we recommend that future research should verify if associations between formant position and self-reported personality traits are truly negligible in size, potentially with an even larger sample size, that might help to find definitive answers to this research question.

4.5. Limitations and further implications for future research

We would also like to note some limitations of the current study. First, the voice recordings we used differed in content and language, potentially reducing comparability. However, it has been reported that perceptions of voice recordings differing in length

and content show moderate to strong correlations (Mahrholz et al., 2018; Puts et al., 2012) and we accounted for variations across datasets in our analyses. Second, different personality questionnaires were used across datasets, potentially leading to more random noise, that we, again, attempted to control in our models. Third, we gathered information about speaker's personality via self-report questionnaires only. Self-report personality questionnaires have sometimes been criticized for their informational and motivational limitations (e.g. Breil et al., 2019; Gosling, John, Craik, & Robins, 1998; John & Robins, 1994). Thus, future studies should investigate whether the associations between voice pitch and personality traits that we report are robust when gathering personality traits via other reports (from friends or family members) or behavioral traces. Fourth, although we included data collected in four different countries (see Table 2), all of our data stem from WEIRD countries (Henrich, Heine, & Norenzayan, 2010), and our results might thus not generalize to other countries and cultures. Fifth, we cannot draw conclusions about how people might alter their voice pitch in different situations to express their personality, as we investigated associations between self-reported personality traits and voice pitch using (semi-)standardized voice recordings that did not include a social context. Hence, future research should investigate whether changes in voice pitch and personality states in different social situations might be associated. Sixth, the ROPE (-0.1; 0.1) we chose means that we called effects negligible, if a change in voice pitch that is approximately 5% of the sex difference in voice pitch would yield less than a standard deviation change in the outcome. In other words, we interpreted effects as substantial where all values in the 89% HDI were bigger than 0.1 or smaller than -0.1, which translates to a rather small effect size. Of course, all ROPEs are somewhat arbitrary. Given the information reported here, other researchers could potentially contrast the estimated associations against the just noticeable difference in voice pitch, but as there are no comparable just noticeable differences for, for instance, extraversion, and as we did not examine perception here, we chose not to do so. Still, all necessary information is available. Seventh, we did not collect perceiver ratings of personality based on the voice recordings and therefore we cannot draw conclusions about whether people use voice pitch as a cue to form social impressions about other people's personality (e.g., Breil, Osterholz, Nestler, & Back, 2021) or whether perceptions based on voice pitch are accurate. Investigating personality perceptions and their accuracy based on voice pitch might be another interesting avenue of future research. Eighth, in contrast to previously published studies, in our data, women self-reported as being more dominant and more unrestricted in sociosexual behavior than men. This effect might be a result of specific sample characteristics (e.g. having only data from women in sample 1) that might be a research question of interest in future studies.

5. Conclusion

In a large multilab project, we investigated the association between measured voice parameters and self-reported personality traits. We provide the first evidence that voice pitch might be a valid cue to at least some self-reported personality traits in men and women, including extraversion, dominance, sociosexual orientation including the facets behavior, attitude, and desire. Thus, personality might at least partly be expressed in voice pitch in men and women. We recommend that future research should investigate whether associations between self-reported personality and voice pitch are due to joint physiological mechanisms, such as the influence of testosterone on gene expression in both vocal tissues and the brain, or the influence of personality on both emotional activation and vocal fold tension. Other directions for

future research might be to investigate whether results are robust across different social situations or when using other-ratings rather than self-reported personality.

References

- Allport, G. W., & Cantril, H. (1934). Judging personality from voice. *The Journal of Social Psychology*, 5, 37–55. <https://doi.org/10.1080/00224545.1934.9921582>.
- Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, 111, 256–274. <https://doi.org/10.1037/0033-2909.111.2.256>.
- Apicella, C. L., Feinberg, D. R., & Marlowe, F. W. (2007). Voice pitch predicts reproductive success in male hunter-gatherers. *Biology Letters*, 3, 682–684. <https://doi.org/10.1098/rsbl.2007.0410>.
- Apple, W., Streeter, L. A., & Krauss, R. M. (1979). Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*, 37, 715. <https://doi.org/10.1037/0022-3514.37.5.715>.
- Archer, J. (2009). Does sexual selection explain human sex differences in aggression? *Behavioral and Brain Sciences*, 32, 249–266. <https://doi.org/10.1017/S0140525X09990951>.
- Archer, J., Graham-Kevan, N., & Davies, M. (2005). Testosterone and aggression: A reanalysis of Book, Starzyk, and Quinsey's (2001) study. *Aggression and Violent Behavior*, 10, 241–261. <https://doi.org/10.1016/j.avb.2004.01.001>.
- Aronovitch, C. D. (1976). The voice of personality: Stereotyped judgments and their relation to voice quality and sex of speaker. *The Journal of Social Psychology*, 99, 207–220. <https://doi.org/10.1080/00224545.1976.9924774>.
- Asendorpf, J. B., Penke, L., & Back, M. D. (2011). From dating to mating and relating: Predictors of initial and long-term outcomes of speed-dating in a community sample. *European Journal of Personality*, 25, 16–30. <https://doi.org/10.1002/per.768>.
- Aung, T., & Puts, D. (2020). Voice pitch: A window into the communication of social power. *Current Opinion in Psychology*, 33, 154–161. <https://doi.org/10.1016/j.copsyc.2019.07.028>.
- Bachorowski, J. A., & Owren, M. J. (1999). Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *The Journal of the Acoustical Society of America*, 106, 1054–1063. <https://doi.org/10.1121/1.427115>.
- Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods*, 40, 531–539. <https://doi.org/10.3758/BRM.40.2.531>.
- Boersma, P., & Weenink, D. (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.37, retrieved 11 October 2018 from <http://www.praat.org/>.
- Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology*, 62, 645. <https://doi.org/10.1037/0022-3514.62.4.645>.
- Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology*, 86, 599. <https://doi.org/10.1037/0022-3514.86.4.599>.
- Borkenau, P., & Ostendorf, F. (1993). *NEO-Fünf-Faktoren Inventar (NEO-FFI) (NEO five-factor inventory)*. Göttingen, Germany: Verlag für Psychologie.
- Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, 82, 55–59. <https://doi.org/10.1016/j.anbehav.2011.03.024>.
- Breil, S. M., Geukes, K., Wilson, R. E., Nestler, S., Vazire, S., & Back, M. D. (2019). Zooming into real-life extraversion—How personality and situation shape sociability in social interactions. *Collabra Psychology*, 5, 1–28. <https://doi.org/10.1525/collabra.170>.
- Breil, S. M., Osterholz, S., Nestler, S., & Back, M. D. (2021). Contributions of nonverbal cues to the accurate judgment of personality traits. In T. D. Letzring & J. S. Spain (Eds.), *The Oxford handbook of accurate personality judgment* (pp. 195–218). Oxford University Press.
- Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80. <https://doi.org/10.18637/jss.v080.i01>.
- Burri, A., Spector, T., & Rahman, Q. (2015). Common genetic factors among sexual orientation, gender nonconformity, and number of sex partners in female twins: Implications for the evolution of homosexuality. *The Journal of Sexual Medicine*, 12, 1004–1011. <https://doi.org/10.1111/jsm.12847>.
- Campbell, A. (2004). Female competition: Causes, constraints, content, and contexts. *Journal of Sex Research*, 41, 16–26. <https://doi.org/10.1080/00224490409552210>.
- Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2015). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76, 1–29. <https://doi.org/10.18637/jss.v076.i01>.
- Collins, S. A. (2000). Men's voices and women's choices. *Animal Behaviour*, 60, 773–780. <https://doi.org/10.1006/anbe.2000.1523>.
- Driebe, J. C., Sidari, M. J., Dufner, M., von der Heiden, J. M., Bürkner, P.-C., Penke, L., Zietsch, B. P., & Arslan, R. C. (2020). Intelligence can be detected but is not found attractive in videos and live interactions. Preprint on PsyArxiv. <https://dx.doi.org/10.31234/osf.io/qv3eh>.
- Fairbanks, G. (1960). *Voice and Articulation Drillbook* (2nd ed.). New York: Harper & Row.
- Feinberg, D. R. (2018, November 6). Praat scripts. <https://doi.org/10.17605/OSF.IO/HUZ7D>.
- Feinberg, D. R., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2008). The role of femininity and averageness of voice pitch in aesthetic judgments of women's voices. *Perception*, 37, 615–623. <https://doi.org/10.1068/p5514>.
- Feinberg, D. R., Jones, B. C., & Armstrong, M. M. (2019). No evidence that men's voice pitch signals formidability. *Trends in Ecology & Evolution*, 34, 190–192. <https://doi.org/10.1016/j.tree.2018.12.014>.
- Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, 69, 561–568. <https://doi.org/10.1016/j.anbehav.2004.06.012>.
- Feinberg, D. R., Jones, B. C., DeBruine, L. M., O'Connor, J. J. M., Tigue, C. C., & Borak, D. J. (2011). Integrating fundamental and formant frequencies in women's preferences for men's voices. *Behavioral Ecology*, 22, 1320–1325. <https://doi.org/10.1093/beheco/arr134>.
- Fisher, C. I., Hahn, A. C., DeBruine, L. M., & Jones, B. C. (2016). Is women's sociosexual orientation related to their physical attractiveness? *Personality and Individual Differences*, 101, 396–399. <https://doi.org/10.1016/j.paid.2016.06.022>.
- Gosling, S. D., John, O. P., Craik, K. H., & Robins, R. W. (1998). Do people know how they behave? Self-reported act frequencies compared with online codings by observers. *Journal of Personality and Social Psychology*, 74, 1337–1349. <https://doi.org/10.1037/0022-3514.74.5.1337>.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD 29–29. *Nature*, 466. <https://doi.org/10.1038/466029a>.
- Hill, A. K., Hunt, J., Welling, L. L., Cárdenas, R. A., Rotella, M. A., Wheatley, J. R., ... Puts, D. A. (2013). Quantifying the strength and form of sexual selection on men's traits. *Evolution and Human Behavior*, 34, 334–341. <https://doi.org/10.1016/j.evolhumbehav.2013.05.004>.
- Huang, G., Pencina, K. M., Coady, J. A., Beleva, Y. M., Bhasin, S., & Basaria, S. (2015). Functional voice testing detects early changes in vocal pitch in women during testosterone administration. *The Journal of Clinical Endocrinology & Metabolism*, 100, 2254–2260. <https://doi.org/10.1210/jc.2015-1669>.
- Hughes, S. M., Dispenza, F., & Gallup, G. G. Jr. (2004). Ratings of voice attractiveness predict sexual behavior and body configuration. *Evolution and Human Behavior*, 25, 295–304. <https://doi.org/10.1016/j.evolhumbehav.2004.06.001>.
- Hodges-Simeon, C. R., Gaulin, S. J., & Puts, D. A. (2010). Different vocal parameters predict perceptions of dominance and attractiveness. *Human Nature*, 21, 406–427. <https://doi.org/10.1007/s12110-010-9101-5>.
- Irwig, M. S., Childs, K., & Hancock, A. B. (2017). Effects of testosterone on the transgender male voice. *Andrology*, 5, 107–112. <https://doi.org/10.1111/andr.12278>.
- Jacobs, I., & Scholl, W. (2005). Interpersonale Adjektivliste (IAL). *Diagnostica*, 51, 145–155. <https://doi.org/10.1026/0012-1924.51.3.145>.
- Jadoul, Y., Thompson, B., & de Boer, B. (2018). Introducing Parselmouth: A Python interface to Praat. *Journal of Phonetics*, 71, 1–15. <https://doi.org/10.1016/j.jwocn.2018.07.001>.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The big five inventory—versions 4a and 5d*. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.
- John, O. P., & Robins, R. W. (1994). Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism. *Journal of Personality and Social Psychology*, 66, 206–219. <https://doi.org/10.1037/0022-3514.66.1.206>.
- Jones, B. C., Hahn, A. C., Fisher, C. I., Wang, H., Kandrik, M., & DeBruine, L. M. (2018). General sexual desire, but not desire for uncommitted sexual relationships, tracks changes in women's hormonal status. *Psychoneuroendocrinology*, 88, 153–157. <https://doi.org/10.1016/j.psychoneuroend.2018.03.007>.
- Jünger, J., Kordsmeyer, T., Gerlach, T. M., & Penke, L. (2018). Fertile women evaluate male bodies as more attractive, regardless of masculinity. *Evolution and Human Behavior*, 39, 412–423. <https://doi.org/10.1016/j.evolhumbehav.2018.03.007>.
- Jünger, J., Motta-Mena, N. V., Cardenas, R., Bailey, D., Rosenfield, K. A., Schild, C., ... Puts, D. A. (2018). Do women's preferences for masculine voices shift across the ovulatory cycle? *Hormones and Behavior*, 106, 122–134. <https://doi.org/10.1016/j.yhbeh.2018.10.008>.
- Klofstad, C. A., Anderson, R. C., & Nowicki, S. (2015). Perceptions of competence, strength, and age influence voters to select leaders with lower-pitched voices. *PLoS One*, 10, e0133779. <https://doi.org/10.1371/journal.pone.0133779>.
- Kordsmeyer, T. L., Hunt, J., Puts, D. A., Ostner, J., & Penke, L. (2018). The relative importance of intra- and intersexual selection on human male sexually dimorphic traits. *Evolution and Human Behavior*, 39, 424–436. <https://doi.org/10.1016/j.evolhumbehav.2018.03.008>.
- Kordsmeyer, T. L., & Penke, L. (2019). Effects of male testosterone and its interaction with cortisol on self- and observer-rated personality states in a competitive mating context. *Journal of Research in Personality*, 78, 76–92. <https://doi.org/10.1016/j.jrp.2018.11.001>.
- Kruschke, J. K. (2014). *Doing bayesian data analysis: A tutorial with r, jags, and stan*. Academic Press.
- Kruschke, J. K. (2018). Rejecting or accepting parameter values in bayesian estimation. *Advances in Methods and Practices in Psychological Science*, 1, 270–280. <https://doi.org/10.1177/2515245918771304>.
- Lang, F. R., Lüdtke, O., & Asendorpf, J. B. (2001). Validity and psychometric equivalence of the German version of the Big Five Inventory in young, middle-aged and old adults. *Diagnostica*, 73, 111–121.
- Leongómez, J. D., Binter, J., Kubíčová, L., Stolařová, P., Klapilová, K., Havlíček, J., & Roberts, S. C. (2014). Vocal modulation during courtship increases proceptivity

- even in naive listeners. *Evolution and Human Behavior*, 35, 489–496 <https://doi.org/10.1016/j.evolhumbehav.2014.06.00>.
- Leongómez, J. D., Mileva, V. R., Little, A. C., & Roberts, S. C. (2017). Perceived differences in social status between speaker and listener affect the speaker's vocal characteristics. *PLoS One*, 12, e0179407. <https://doi.org/10.1371/journal.pone.0179407>.
- Mahrholz, G., Belin, P., & McAleer, P. (2018). Judgements of a speaker's personality are correlated across differing content and stimulus type. *PLoS One*, 13, e0204991. <https://doi.org/10.1371/journal.pone.0204991>.
- Malloy, E. B., & Miller, V. R. (1958). A possible basis for the association of voice characteristics and personality traits. *Communications Monographs*, 25, 255–260. <https://doi.org/10.1080/03637755809375240>.
- Maikowski, D., & Lüdecke, D. (2019). Understand and describe bayesian models and posterior distributions using BayestestR. CRAN. <https://doi.org/10.5281/zenodo.2556486>.
- Mayew, W. J., Parsons, C. A., & Venkatachalam, M. (2013). Voice pitch and the labor market success of male chief executive officers. *Evolution and Human Behavior*, 34, 243–248. <https://doi.org/10.1016/j.evolhumbehav.2013.03.001>.
- McAleer, P., Todorov, A., & Belin, P. (2014). How do you say 'Hello'? Personality impressions from brief novel voices. *PLoS One*, 9, e90779. <https://doi.org/10.1371/journal.pone.0090779>.
- McCrae, R. R., & Costa, P. T. (1989). The structure of interpersonal traits: Wiggins's circumplex and the five-factor model. *Journal of Personality and Social Psychology*, 56, 586–595. <https://doi.org/10.1037/0022-3514.56.4.586>.
- McElreath, R. (2018). *Statistical rethinking: A bayesian course with examples in r and stan*. Chapman: Hall/CRC.
- Mileva, M., Tompkinson, J., Watt, D., & Burton, A. M. (2018). Audiovisual integration in social evaluation. *Journal of Experimental Psychology*, 44, 128–138. <https://doi.org/10.1037/xhp0000439>.
- Naumann, L. P., Vazire, S., Rentfrow, P. J., & Gosling, S. D. (2009). Personality judgments based on physical appearance. *Personality and Social Psychology Bulletin*, 35, 1661–1671. <https://doi.org/10.1177/0146167209346309>.
- Ostendorf, F. (2001). Measuring interpersonal behavior with the German interpersonal adjective scales. In *Personality and temperament: Genetics, evolution, and structure* (pp. 232–260). Lengerich: Pabst Scientific Publishers.
- Ostendorf, F., & Angleitner, A. (2003). *NEO-Persönlichkeitsinventar (revidierte Form, NEO-PI-R) nach Costa und McCrae*. Göttingen: Hogrefe.
- Ostovich, J. M., & Sabini, J. (2004). How are sociosexuality, sex drive, and lifetime number of sexual partners related? *Personality and Social Psychology Bulletin*, 30, 1255–1266. <https://doi.org/10.1177/0146167204264754>.
- Penke, L., & Asendorpf, J. B. (2008). Beyond global sociosexual orientations: A more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of Personality and Social Psychology*, 95, 1113–1135. <https://doi.org/10.1037/0022-3514.95.5.1113>.
- Puts, D. A. (2005). Mating context and menstrual phase affect women's preferences for male voice pitch. *Evolution and Human Behavior*, 26, 388–397. <https://doi.org/10.1016/j.evolhumbehav.2005.03.001>.
- Puts, D. A. (2006). Cyclic variation in women's preferences for masculine traits: Potential hormonal causes. *Human Nature*, 17, 114–127. <https://doi.org/10.1007/s12110-006-1023-x>.
- Puts, D. A., Apicella, C. L., & Cárdenas, R. A. (2012). Masculine voices signal men's threat potential in forager and industrial societies. *Proceedings of the Royal Society B: Biological Sciences*, 279, 601–609. <https://doi.org/10.1098/rspb.2011.0829>.
- Puts, D. A., & Aung, T. (2019). Does men's voice pitch signal formidability? A reply to Feinberg et al. *Trends in Ecology & Evolution*, 34, 189–190. <https://doi.org/10.1016/j.tree.2018.12.004>.
- Puts, D. A., & Cárdenas, R. (2018, December 3). Voice scripts. <https://doi.org/10.17605/OSF.IO/K2BHS>.
- Puts, D. A., Doll, L. M., & Hill, A. K. (2014). Sexual selection on human voices. In *Evolutionary perspectives on human sexual psychology and behavior* (pp. 69–86). New York, NY: Springer.
- Puts, D. A., Gaulin, S. J., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, 27, 283–296. <https://doi.org/10.1016/j.evolhumbehav.2005.11.003>.
- Puts, D. A., Hill, A. K., Bailey, D. H., Walker, R. S., Rendall, D., Wheatley, J. R., ... Ramos-Fernandez, G. (2016). Sexual selection on male vocal fundamental frequency in humans and other anthropoids. *Proceedings of the Royal Society B: Biological Sciences*, 283. <https://doi.org/10.1098/rspb.2015.2830>.
- Puts, D. A., Hodges, C. R., Cárdenas, R. A., & Gaulin, S. J. (2007). Men's voices as dominance signals: Vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior*, 28, 340–344. <https://doi.org/10.1016/j.evolhumbehav.2007.05.002>.
- Puts, D. A., Pope, L. E., Hill, A. K., Cárdenas, R. A., Welling, L. L., Wheatley, J. R., & Breedlove, S. M. (2015). Fulfilling desire: Evidence for negative feedback between men's testosterone, sociosexual psychology, and sexual partner number. *Hormones and Behavior*, 70, 14–21. <https://doi.org/10.1016/j.yhbeh.2015.01.006>.
- Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41, 203–212. <https://doi.org/10.1016/j.jrp.2006.02.001>.
- Rosenfield, K. A., Sorokowska, A., Sorokowski, P., & Puts, D. A. (2020). Sexual selection for low male voice pitch among Amazonian forager-horticulturists. *Evolution and Human Behavior*, 41, 3–11. <https://doi.org/10.1016/j.evolhumbehav.2019.07.002>.
- R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Scherer, K. R. (1972). Judging personality from voice: A cross-cultural approach to an old issue in interpersonal perception 1. *Journal of Personality*, 40, 191–210. <https://doi.org/10.1111/j.1467-6494.1972.tb00998.x>.
- Scherer, K. R. (1978). Personality inference from voice quality: The loud voice of extroversion. *European Journal of Social Psychology*, 8, 467–487. <https://doi.org/10.1002/ejsp.2420080405>.
- Schild, C., Stern, J., & Zettler, I. (2020). Linking men's voice pitch to actual and perceived trustworthiness across contexts. *Behavioral Ecology*, 31, 164–175. <https://doi.org/10.1093/beheco/arz173>.
- Schild, C., Stern, J., Penke, L., & Zettler, I. (in press). Voice pitch – a valid indicator of one's unfaithfulness in committed relationships? *Adaptive Human Behavior and Physiology*. <https://doi.org/10.1007/s40750-020-00154-0>.
- Sell, A., Bryant, G. A., Cosmides, L., Tooby, J., Sznycer, D., Von Rueden, C., ... Gurven, M. (2010). Adaptations in humans for assessing physical strength from the voice. *Proceedings of the Royal Society B*, 276, 575–584. <https://doi.org/10.1098/rspb.2010.0769>.
- Sellers, J. G., Mehl, M. R., & Josephs, R. A. (2007). Hormones and personality: Testosterone as a marker of individual differences. *Journal of Research in Personality*, 41, 126–138. <https://doi.org/10.1016/j.jrp.2006.02.004>.
- Shirazi, T. N., Self, H., Dawood, K., Rosenfield, K. A., Penke, L., Carré, J. M., ... Puts, D. A. (2019). Hormonal predictors of women's sexual motivation. *Evolution and Human Behavior*, 40, 336–344. <https://doi.org/10.1016/j.evolhumbehav.2019.02.002>.
- Stern, J., Karastoyanova, K., Kandrik, M., Torrance, J., Hahn, A. C., Holzleitner, I., ... Jones, B. C. (2020). Are sexual desire and sociosexual orientation related to men's salivary steroid hormones? *Adaptive Human Behavior and Physiology*, 6, 447–466. <https://doi.org/10.1007/s40750-020-00148-y>.
- Stern, J., Kordsmeyer, T. L., & Penke, L. (2021). A longitudinal evaluation of ovulatory cycle shifts in women's mate attraction and preferences. *Hormones and Behavior*, 128, 104916. <https://doi.org/10.1016/j.yhbeh.2020.104916>.
- Stillman, T. F., & Maner, J. K. (2009). A sharp eye for her SOI: Perception and misperception of female sociosexuality at zero acquaintance. *Evolution and Human Behavior*, 30, 124–130. <https://doi.org/10.1016/j.evolhumbehav.2008.09.005>.
- Tigue, C. C., Borak, D. J., O'Connor, J. J., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. *Evolution and Human Behavior*, 33, 210–216. <https://doi.org/10.1016/j.evolhumbehav.2011.09.004>.
- Vehtari, A., Gelman, A., & Gabry, J. (2016). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*. <https://doi.org/10.1007/s11222-016-9696-4>.
- Wood, S. N., Pya, N., & Säfken, B. (2016). Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association*, 111, 1548–1563. <https://doi.org/10.1080/01621459.2016.1180986>.
- Xu, Y., Lee, A., Wu, W. L., Liu, X., & Birkholz, P. (2013). Human vocal attractiveness as signaled by body size projection. *PLoS One*, 8, e62397. <https://doi.org/10.1371/journal.pone.0062397>.
- Zuckerman, M., & Driver, R. E. (1989). What sounds beautiful is good: The vocal attractiveness stereotype. *Journal of Nonverbal Behavior*, 13, 67–82. <https://doi.org/10.1007/BF00990791>.
- Zuckerman, M., Miyake, K., & Elkin, C. S. (1995). Effects of attractiveness and maturity of face and voice on interpersonal impressions. *Journal of Research in Personality*, 29, 253–272. <https://doi.org/10.1006/jrpe.1995.1015>.