Supplemental Material - Part A "Big Five Personality Perceptions from Voices and Faces – Impressions and Kernels of Truth"

author1 author2 author3 author4 and author5

27th of October 2023

Contents

1 Introduction

This supplementary material comprises a substantial collection of additional information pertaining to two voice experiments (E1, E2) and one face experiment. In each trial, participants' task was to discriminate between pairs of high and low trait-average voices (E1, E2) or high and low trait-composite faces (E3), as provided by (Kramer and Ward 2010), in accordance with a discrimination statement. The notable distinction between E1 and E2 lies in the source of the stimuli utilized. In the case of E1, the stimuli were morphed through averaging voices based on speakers' self-ratings, while for E2, the foundation for stimuli creation relied on listeners' impressions of individual voice recordings of the speakers.

2 Details on Speaker Selection for E1 and E2

2.1 Speaker Selection for Experiment E1 (Trait-Averages Based on Self-Ratings)

For the online experiment E1, we selected for each of the Big Five traits and per speaker sex, the five speakers with the lowest and highest **self-rated** trait values in order to generate low and high trait-average voices for each speaker sex and speaker trait. Whenever there were two or more speakers with the same trait value, we selected the speaker, whose mean absolute distance to the median of the other four trait values was lowest. Panel **A** of Figure A1 shows the self-rated trait values and panel **B** shows the respective mean listeners' evaluations of the selected female and male speakers for the high and low trait speakers, as well as speakers, who were not included in the respective trait-average.

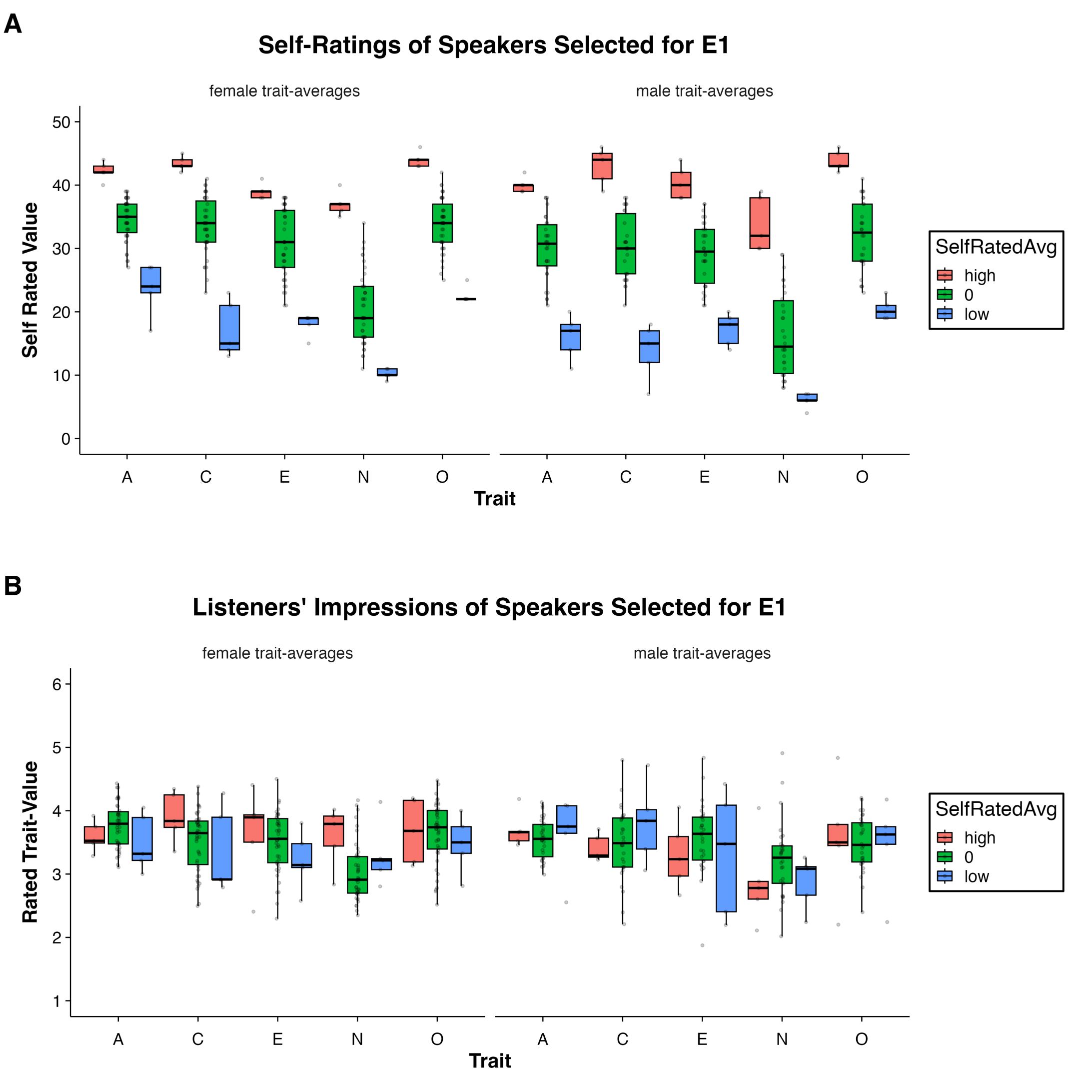

Figure A 1: Panel A illustrates the self-rated trait values and panel B illustrates the mean ratings of the female (left side) and male (right side) speakers that were used as the foundation for the stimuli in Experiment E1. Note that speakers who were morphed into high trait-averages are represented by the color red, speakers who were morphed into low trait-averages are represented by the color blue, and speakers who were not included in the trait-averages are represented by the color green.

2.2 Speaker Selection for Experiment E2 (Trait-Averages based on Listeners' Impressions)

For the online experiment E2, we selected for each of the Big Five traits, the five speakers per speaker sex with the lowest and highest **perceived** trait values in order to generate low and high trait-averages for each speaker sex and speaker trait. Whenever there were two or more speakers with the same trait value, we selected the speaker, whose mean absolute distance to the median of the other four trait values was lowest. Panel **A** of Figure A2 shows the mean listeners' ratings and panel **B** shows the respective self-ratings of the selected female and male speakers for the high and low trait speakers, as well as speakers, who were not included in the respective trait-average.

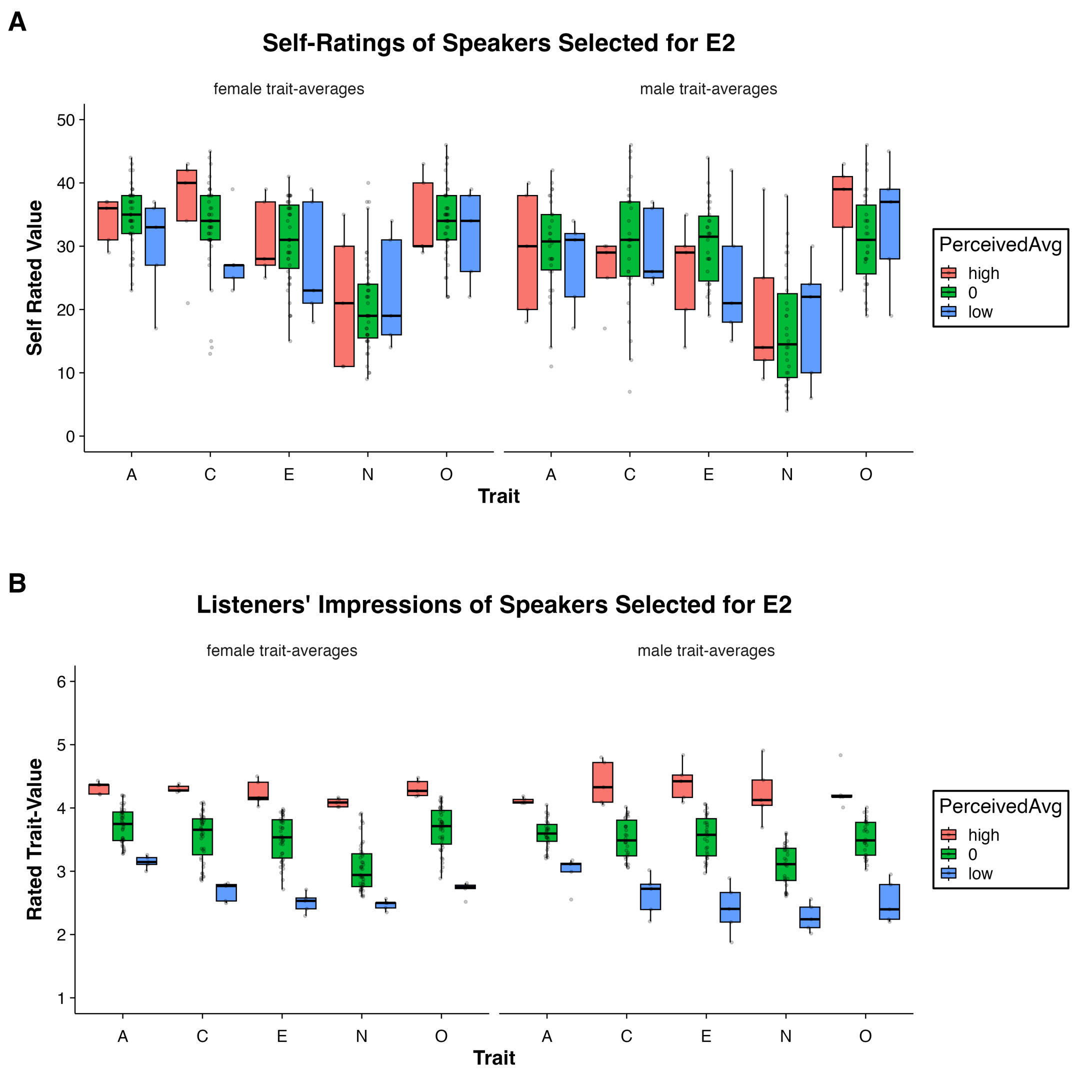

Figure A 2: The box plots illustrate the mean listeners' ratings of the female (f) and male (m) speakers that were utilized as the foundation for the stimuli in Experiment E2. Note that speakers who were morphed into high trait-averages are represented by the color red, speakers who were morphed into low trait-averages are represented by the color blue, and speakers who were not included in the trait-averages are represented by the color green.

3 Details on the Discrimination Statements used in E1, E2, and E3

3.1 Big Five

The following positve and negative discrimination statements for the Big Five low and high trait-average voices were taken and adapted from the Big Five-Inventory-10 (Rammstedt et al. 2002).

Agreeableness

- Welche/r Sprecher/in scheint anderen eher leicht Vertrauen zu schenken und eher an das Gute im Menschen zu glauben? („Which speaker is more likely to be generally trusting?")
- Welche/r Sprecher/in scheint eher dazu zu neigen, andere zu kritisieren? („Which speaker is more likely to find fault with others?")

Conscientiousness

- Welche/r Sprecher/in erledigt Aufgaben eher gründlicher? („Which speaker is more likely to do a thorough job?")
- Welche/r Sprecher/in ist bequemer und neigt eher zur Faulheit? („Which speaker is more likely to be lazy?")

Extraversion

- Welche/r Sprecher/in geht eher aus sich heraus und ist geselliger? („Which speaker is more likely to be outgoing and sociable?")
- Welche/r Sprecher/in ist eher zurückhaltender, reservierter? ("Which speaker is more likely to be reserved?")

Neuroticism

- Welche/r Sprecher/in wird leichter nervös und eher unsicher? („Which speaker is more likely to get nervous easily?")
- Welche/r Sprecher/in ist eher entspannt und durch Stress nicht aus der Ruhe zu bringen? („Which speaker is more likely to be relaxed and handle stress well?")

Openness

- Welche/r Sprecher/in hat eher eine aktivere Vorstellungskraft und ist fantasievoller? („Which speaker has a more active imagination?")
- Welche/r Sprecher/in hat eher weniger künstlerisches Interesse? („Which speaker is more likely to have fewer artistic interests?")

3.2 Autistic Traits (only E1, not analysed)

For the autistic traits we chose items from (Baron-Cohen et al. 2001). The two items, that had the highest loadings on each of the factors AQ~social interaction~ and $\text{AQ}_{\text{attention-to-detail}}$ in (Hoekstra et al. 2008) were adapted to be used as discrimination statements in the voice experiment (AQ~social interaction~: "I find it hard to make new friends." r = .643, "I enjoy doing things spontaneously." r = .755; $\text{AQ}_{\text{attention-to-detail}}$: "I am fascinated by numbers." r = .920, "I am fascinated by dates." r = .754). The item "I am fascinated by dates." was then adapted to load negatively on the $\text{AQ}_{\text{attention-to-detail}}$ factor, so one item had a positive and one a negative loading on each of the two AQ traits. For our German version, we used the translations by (Freitag et al. 2007). The discrimination statements for the autistic traits were

AQ Social Interaction

- Welche/r Sprecher/in fällt es eher schwer, neue Freunde kennen zu lernen? („Which speaker finds it rather hard to make new friends?")
- Welche/r Sprecher/in ist eher gerne spontan? ("Which speaker enjoys doing things spontaneously?")

AQ Attention to Detail

- Welche/r Sprecher/in ist eher von Zahlen fasziniert? („Which speaker is more fascinated by numbers?")
- Welche/r Sprecher/in ist eher nicht von Datumsangaben fasziniert? („Which speaker is less likely to be fascinated by dates and data?")

4 Statistics of the ANOVAs to Analyse Effects of Participant Sex for E1 and E2

Figure A3 shows the results of the ten 2 x 2 x 2 ANOVAs with factors Speaker Sex (SpSex: female, male), Sentence (S1, S2) and Participants Sex (PSex: female, male) on the 'proportion of correct responses', separately per trait for both experiments E1 (top rows) and E2 (bottom rows). Figure A4 and A5 show the visualisations of the 4-way interaction of PSex x SpSex x Sentence x Trait of experiments E1 and E2, respectively.

	PSex			**SpSex**			**Sentence**		
Experiment E1	***F*(1,149)**	***p***	***np2***	***F*(1,149)**	***p***	***np2***	***F*(1,149)**	***p***	***np2***
Agreeableness	1.948	.165	.013	1.358	.246	.009	5.845	.017 *	.038
Consicientiousness	0.074	.785	.001	0.007	.932	.001	17.324	< .001 ***	.104
Extraversion	0.446	.506	.003	310.232	< .001 ***	.676	1.132	.289	.008
Neuroticism	0.446	.505	.003	73.620	.001 ***	.331	2.511	.115	.017
Openness	5.791	.017 *	.037	2.275	.134	.015	21.786	.001 ***	.128
Experiment E2	***F*(1,151)**	***p***	***np2***	***F*(1,151)**	***p***	***np2***	***F*(1,151)**	***p***	***np2***
Agreeableness	0.018	.892	< .001	41.043	< .001 ***	.214	7.444	.007 **	.047
Consicientiousness	2.182	.142	.014	84.471	< .001 ***	.359	32.632	< .001 ***	.178
Extraversion	2.238	.137	.015	30.292	< .001 ***	.167	1.943	.165	.013
Neuroticism	0.874	.351	.006	67.373	< .001 ***	.309	0.775	.380	.005
Openness	0.496	.482	< .001	1.744	.189	.011	0.460	.498	< .001

	PSex x SpSex			**PSex x Sentence**			**SpSex x Sentence**			**PSex x SpSex x Sentence**		
Experment E1	***F*(1,149)**	***p***	***np2***	***F*(1,149)**	***p***	***np2***	***F*(1,149)**	***p***	***np2***	***F*(1,149)**	***p***	***np2***
Agreeableness	1.098	.296	.007	0.133	.715	.001	21.437	< .001 ***	.126	0.268	.605	.002
Consicientiousness	1.525	.219	.010	2.04	.155	.014	0.221	.639	.001	0.003	.958	.001
Extraversion	7.626	.006 **	.049	0.003	.957	.001	1.517	.220	.010	2.927	.089	.019
Neuroticism	0.844	.360	.006	1.184	.278	.008	0.033	.856	.001	0.404	.526	.003
Openness	0.042	.839	.001	5.952	.016	.038	35.302	.001 ***	.192	0.206	.651	.001
Experment E2	***F*(1,151)**	***p***	***np2***	***F*(1,151)**	***p***	***np2***	***F*(1,151)**	***p***	***np2***	***F*(1,151)**	***p***	***np2***
Agreeableness	1.987	.161	.013	0.050	.823	< .001	1.903	.170	.012	0.249	.619	< .001
Consicientiousness	2.956	.088	.019	0.435	.510	.003	18.705	< .001 ***	.110	2.862	.093	.019
Extraversion	6.101	.015 *	.039	0.178	.674	.001	0.704	.403	.005	0.813	.369	.005
Neuroticism	0.147	.702	< .001	0.484	.488	< .001	1.687	.196	.011	0.327	.568	< .001
Openness	1.675	.198	.011	4.902	.028 *	.031	0.355	.552	< .001	0.132	.717	< .001

Figure A 3: Summary of the main effects and the interaction of the 2 x 2 x 2 ANOVAs with factors Speaker Sex (SpSex: female, male), Sentence (S1, S2) and Participants Sex (PSex: female, male) on the 'proportion of correct responses', separately for Experiment E1 (top rows) and E2 (bottom rows). Note that data of the diverse participants was not included in these analyses.

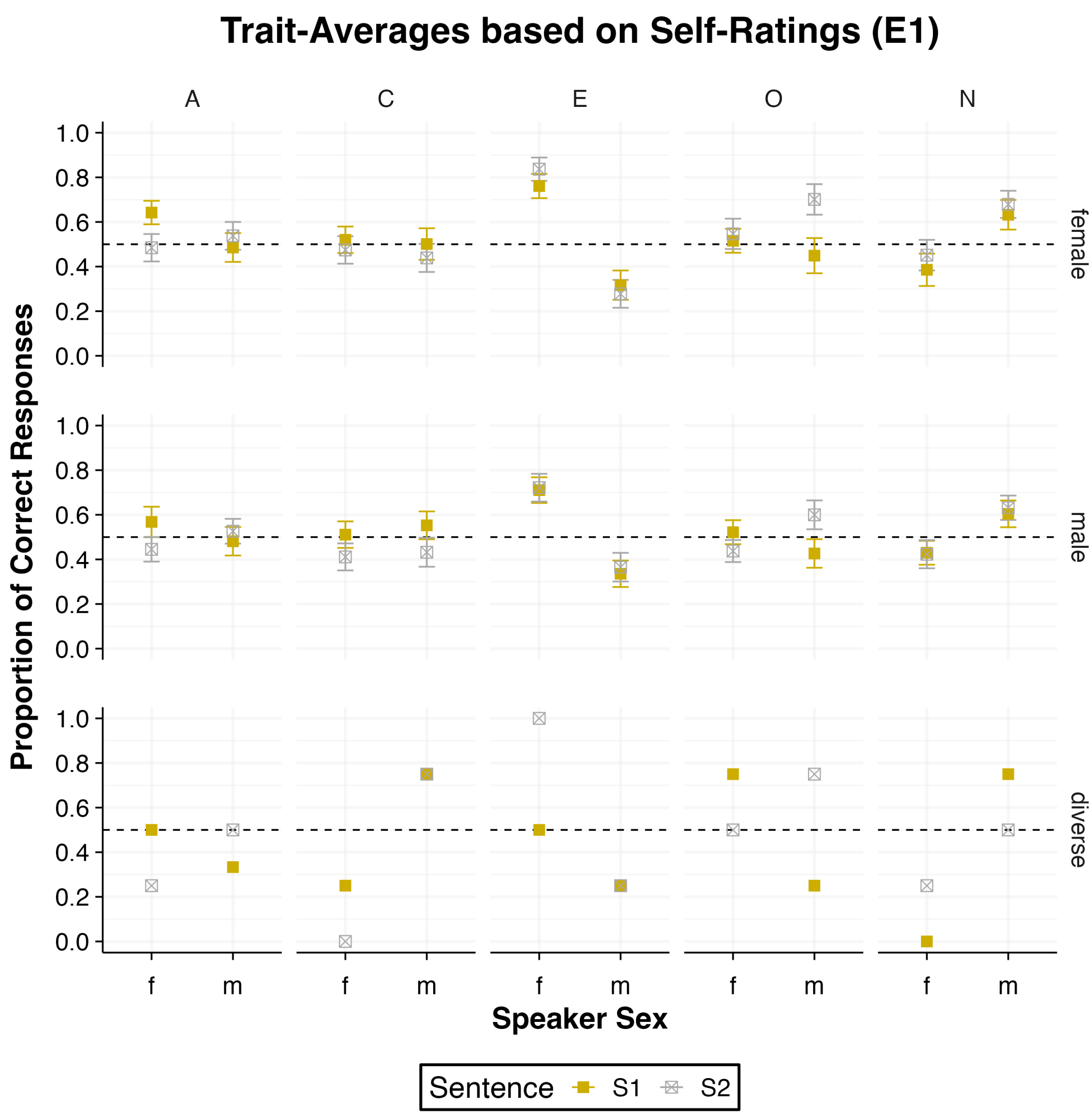

Figure A 4: Experiment E1 - Visualisation of the 'proportion of correct responses' (M and 95%CI) for trait discrimination of high and low trait-averages based on speakers' self-ratings. Visualisation of the 4-way interaction of Participant Sex (PSex: diverse, female, male), Speaker Sex (SpSex: female, male), Sentence (S1, S2) and Trait (A: Agreeableness, C: Conscientiousness, E: Extraversion, N: Neuroticism, and O: Openness).

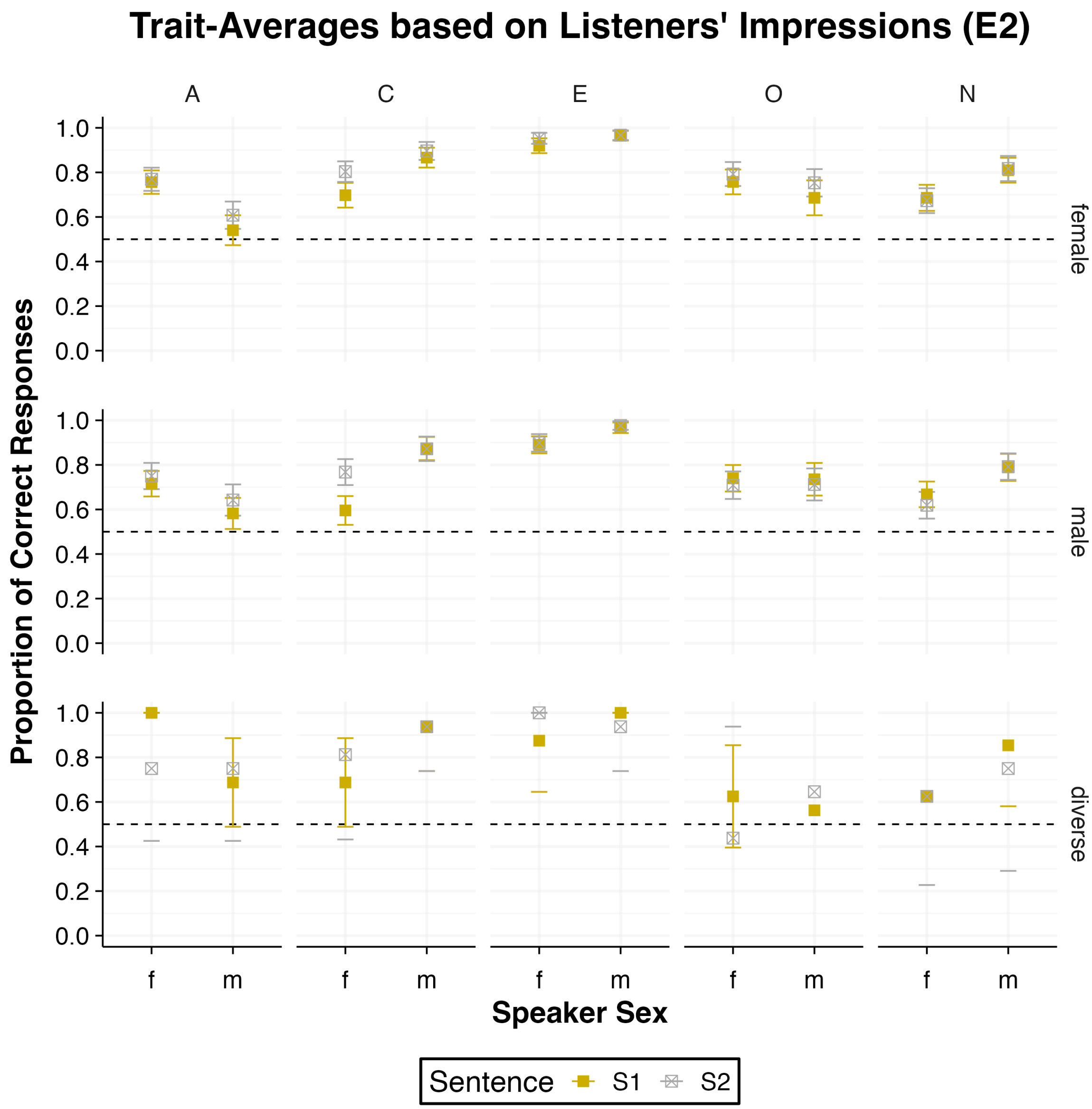

Figure A 5: Experiment E2 - Visualisation of the 'proportion of correct responses' (M and 95%CI) for trait discrimination of high and low trait-averages based on listeners' impressions. Visualisation of the 4-way interaction of Participant Sex (PSex: diverse, female, male), Speaker Sex (SpSex: female, male), Sentence (S1, S2) and Trait (A: Agreeableness, C: Conscientiousness, E: Extraversion, N: Neuroticism, and O: Openness).

5 Correlations Between Vocal and Facial Trait Discrimination Ability Separately per Participant Sex and Trait

Separately per experiment and participant sex, we correlated participants' abilities (Pearson, one-tailed) to discriminate either self-rated (E1) or stereotyped (E2) vocal traits with their ability to discriminate self-rated facial traits (E3). As can be seen in Figure A6, we found for E1 a correlations for male (r = .260, p = .012, N = 76) but not female participants (r = .049, p = .338, N = 75). The opposite pattern was found for E2, where female participants exhibited a positive correlation (r = .270, p = .009, N = 78) but not male participants (r = .015, p = .449, N = 75).

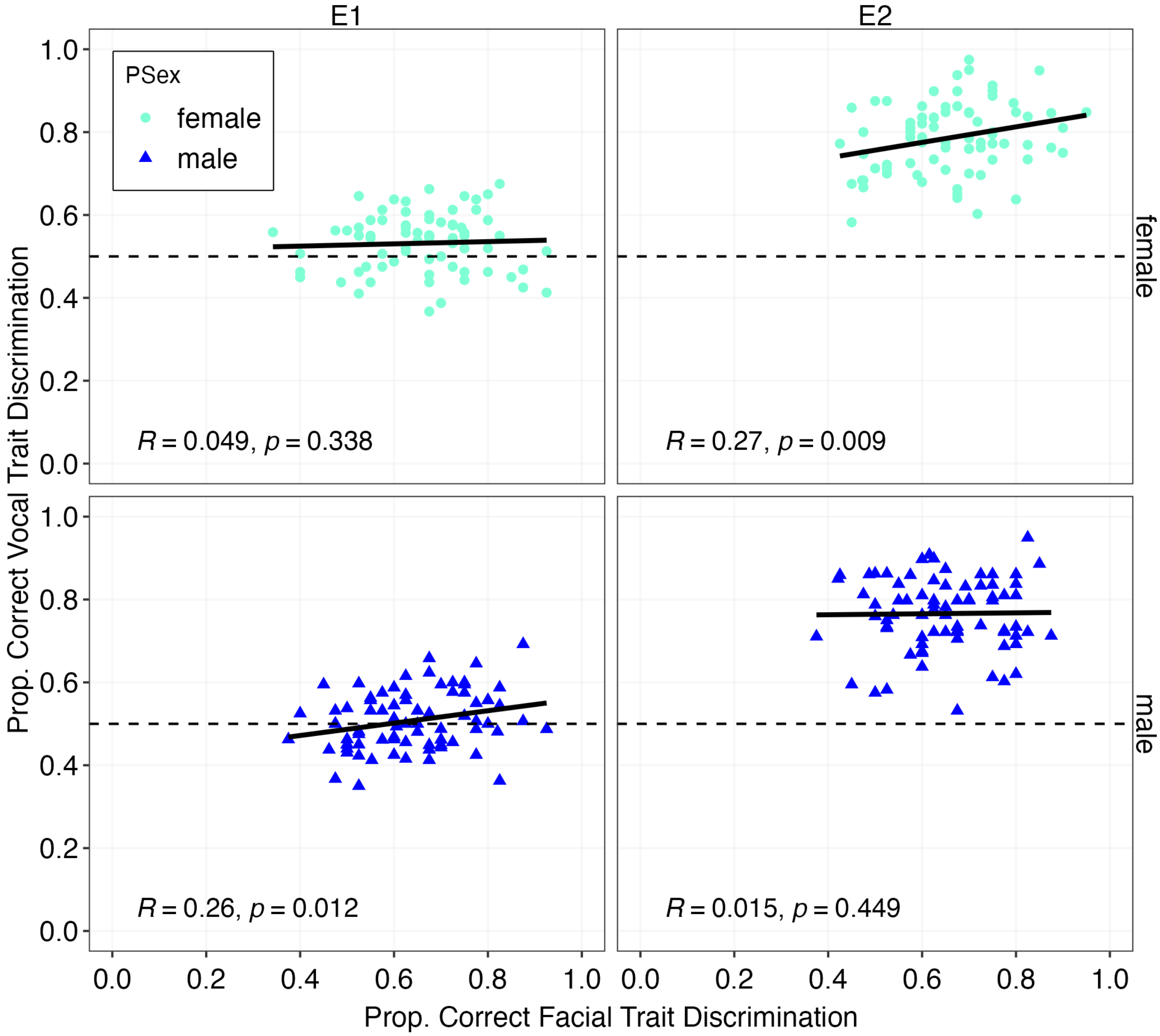

Figure A 6: Pearson, (one-tailed) correlations, between participants' abilities to discriminate either self-rated (E1) or stereotyped (E2) vocal traits with their ability to discriminate self-rated facial traits (E3), separately per experiment (E1, E2) and participant sex (PSex: female, male).

When examining correlations separately per experiment, trait and participant sex (cf. Figure A7), we found that the correlation for self-rated extraversion (E1) was significant for male (r = .240, p = .020, N = 76) but not female (r = .099, p = .198, N = 75) participants. A similar pattern was observed for stereotyped vocal (E2) and self-rated facial extraversion. This correlation was driven by male participants (r = .250, p = .014,

$N = 75$), and not by female participants who performed very well throughout ($r < .01$, $p >. 5$, $N = 78$).

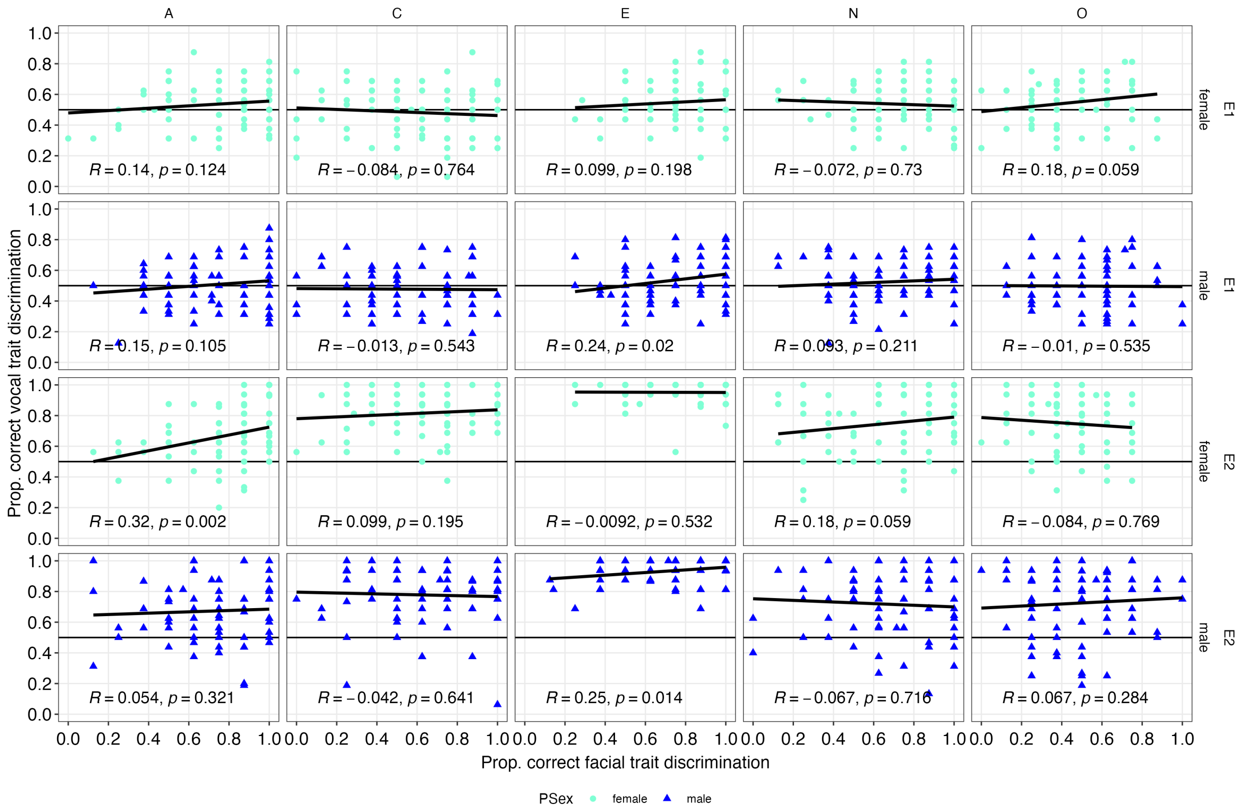

Figure A 7: Pearson, (one-tailed) correlations, between participants' abilities to discriminate either self-rated (E1) or stereotyped (E2) vocal traits with their ability to discriminate self-rated facial traits (E3), separately per experiment (E1, E2), Participant Sex (PSex: female, male) and Trait (A: Agreeableness, C: Conscientiousness, E: Extraversion, N: Neuroticism, and O: Openness).

6 Correlations Between Autistic Traits and Vocal and Facial Trait Discrimination Ability

For males, the mean self-rated AQ score was 19.23 (SD = 6.45, N = 151), for females it was 17.18 (SD = 6.12, N = 153), and for diverse participants, it was 17.60 (SD = 8.56, N = 5). Although we investigated the hypothesis that individuals with higher AQ scores would experience more difficulties in accurately perceiving traits, we calculated Pearson correlations two-tailed, and separately assessed effects of AQattention and AQsocial. This was because these AQ subscores were sometimes reported to show opposite relationships with social perception skills for men and women (Davis et al. 2017), (Rhodes et al. 2013).

6.1 AQ and Vocal Trait Discrimination Ability

There was no correlation between participants' AQ and vocal trait discrimination, neither for self-rated trait-averages (r = -.093, p = .254, N = 152) nor for trait-averages based on listeners' impressions (r = .130, p = .109, N = 157). As a qualification and in contrast to our expectations, male participants with a higher AQ were better in discriminating stereotyped stimuli (r = .290, p = .012, N = 75), an effect that tended to be driven by the AQsocial factor (r = .260, p = .026, N = 75), though with potential contribution of the AQattention factor (r = .150, p = .205, N = 75).

6.2 AQ and Facial Trait Discrimination Ability

There was also no significant correlation between participants' AQ and their facial trait discrimination ability, r = -.090, p = .113, N = 309. As a qualification and in line with our expectations, female participants with a higher AQ performed less well in discriminating facial self-rated stimuli (r = -.170, p = .037, N = 153) , an effect that could not be clearly attributed to the AQsocial (r = -.140, p = .081, N = 153), or the AQattention factor (r = -.130, p = .100, N = 153). There were no such correlations for male participants (p > .3, N = 151).

6.3 Discussion of these results

To the best of our knowledge, our study also represents the first exploration of a potential link between participants' autistic traits, as measured by the AQ (Baron-Cohen et al. 2001; Freitag et al. 2007) and their ability to perceive personality from trait-average voices or faces. We found no evidence for a negative correlation between the AQ (or its sub-scores; Hoekstra et al. (2008)) and participants' performance in personality perception for either voices or faces. This may represent an interesting contrast to other recent findings that individuals with lower autistic traits perform better in emotional expression recognition (Hosokawa et al. 2015; Lewis, Lefevre, and Young 2016; Poljac, Poljac, and Wagemans 2013), although this association is not consistently found (Halliday et al. 2014; Palermo et al. 2018).

Although the present data did not reveal any consistent associations between personality perception and autistic traits, this was qualified by a positive association in females, for whom higher AQsocial traits were unexpectedly correlated with higher performance in facial personality perception. This finding could provide an interesting contrast to recent reports of a negative correlation between a communication-related sub-score of the AQ and performance in the short form of the Geneva Emotion Recognition Test (Ola and Gullon-Scott 2020), as well as to further reports of negative associations between social-cluster autistic traits and face recognition (Davis et al. 2017) and voice identity processing (Schelinski, Borowiak, and Kriegstein 2016; Schelinski, Roswandowitz, and Kriegstein 2017; Skuk et al. 2019). In the absence of specifically powered replications of the present exploratory findings, we caution against an interpretation of these findings at this point in time.

REFERENCES

Baron-Cohen, Simon, Sally Wheelwright, Richard Skinner, Joanne Martin, and Emma Clubley (2001). "The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians”. In: *Journal of autism and developmental disorders* 31.1, pp. 5–17. DOI: 10.1023/a:1005653411471.

Davis, J., E. McKone, M. Zirnsak, T. Moore, R. O’Kearney, D. Apthorp, and R. Palermo (2017). “Social and attention-to-detail sub-clusters of autistic traits differentially predict looking at eyes and face identity recognition ability”. In: *British Journal of Psychology* 108.1, pp. 191–219. DOI: 10.1111/bjop.12188.

Freitag, Christine M, Petra Retz-Junginger, Wolfgang Retz, Christiane Seitz, Haukur Palmason, Jobst Meyer, Michael Rösler, and Alexander von Gontard (2007). “Evaluation der deutschen Version des Autismus-Spektrum-Quotienten (AQ)-die Kurzversion AQ-k”. In: *Zeitschrift für klinische Psychologie und Psychotherapie* 36.4, pp. 280–289. ISSN: 1616-3443. DOI: 10.1026/1616-3443.36.4.280.

Halliday, D. W. R., S. W. S. MacDonald, S. K. Sherf, and J. W. Tanaka (2014). “A reciprocal model of face recognition and autistic traits: Evidence from an individual differences perspective”. In: *Plos One* 9.5. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0133092. URL: %3CGo%20to%20ISI%3E://WOS:000338323200001.

Hoekstra, Rosa A, Meike Bartels, Danielle C Cath, and Dorret I Boomsma (2008). “Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): A study in Dutch population and patient groups”. In: *Journal of autism and developmental disorders* 38, pp. 1555–1566. DOI: 10.1186/2040-2392-5-29.

Hosokawa, M., Y. Nakadoi, Y. Watanabe, S. Sumitani, and T. Ohmori (2015). “Association of autism tendency and hemodynamic changes in the prefrontal cortex during facial expression stimuli measured by multi-channel near-infrared spectroscopy”. In: *Psychiatry and Clinical Neurosciences* 69.3, pp. 145–152. ISSN: 1323-1316. DOI: 10.1111/pcn.12240. URL: %3CGo%20to%20ISI%3E://WOS:000351046600004.

Kramer, Robin S. S. and Robert Ward (2010). “Internal facial features are signals of personality and health”. In: *Quarterly Journal of Experimental Psychology* 63.11, pp. 2273–2287. DOI: 10.1080/17470211003770912.

Lewis, G. J., C. E. Lefevre, and A. W. Young (2016). “Functional architecture of visual emotion recognition ability: a latent variable approach”. In: *Journal of Experimental Psychology-General* 145.5, pp. 589–602. ISSN: 0096-3445. DOI: 10.1037/xge0000160. URL: %3CGo%20to%20ISI%3E://WOS:000374770900006.

Ola, L. and F. Gullon-Scott (2020). “Facial emotion recognition in autistic adult females correlates with alexithymia, not autism”. In: *Autism* 24.8, pp. 2021–2034. ISSN: 1362-3613. DOI: 10.1177/1362361320932727. URL: %3CGo%20to%20ISI%3E://WOS:000551151900001.

Palermo, R., L. Jeffery, J. Lewandowsky, C. Fiorentini, J. L. Irons, A. Dawel, N. Burton, E. McKone, and G. Rhodes (2018). “Adaptive face coding contributes to individual differences in facial expression recognition independently of affective factors”. In: *Journal of Experimental Psychology-Human Perception and Performance* 44.4, pp. 503–517. ISSN: 0096-1523. DOI: 10.1037/xhp0000463. URL: %3CGo%20to%20ISI%3E://WOS:000429682500001.

Poljac, E., E. Poljac, and J. Wagemans (2013). “Reduced accuracy and sensitivity in the perception of emotional facial expressions in individuals with high autism spectrum traits”. In: *Autism* 17.6, pp. 668–680. ISSN: 1362-3613. DOI: 10.1177/1362361312455703. URL: %3CGo%20to%20ISI%3E://WOS:000326151100004.

Rammstedt, Beatrice, Daniel Danner, Christopher J Soto, and Oliver P John (2002). “Validation of the short and extra-short forms of the Big Five Inventory-2 (BFI-2) and their German adaptations”. In: *European Journal of Psychological Assessment* 36.1. ISSN: 2151-2426. DOI: 10.1027/1015-5759/a000481.

Rhodes, G., L. Jeffery, L. Taylor, and L. Ewing (2013). “Autistic traits are linked to reduced adaptive coding of face identity and selectively poorer face recognition in men but not women”. In: *Neuropsychologia* 51.13. Rhodes, Gillian Jeffery, Linda Taylor, Libby Ewing, Louise Ewing, Louise/H-9158-2014, pp. 2702–2708. ISSN: 0028-3932. DOI: 10.1016/Lneuropsychologia.2013.08.016.

Schelinski, S., K. Borowiak, and K. von Kriegstein (2016). “Temporal voice areas exist in autism spectrum disorder but are dysfunctional for voice identity recognition”. In: *Social Cognitive and Affective Neuroscience* 11.11, pp. 1812–1822. ISSN: 1749-5016. DOI: 10.1093/scan/nsw089. URL: %3CGo%20to%20ISI%3E://WOS:000388940000014.

Schelinski, S., C. Roswandowitz, and K. von Kriegstein (2017). “Voice identity processing in autism spectrum disorder”. In: *Autism Research* 10.1, pp. 155–168. ISSN: 1939-3792. DOI: 10.1002/aur.1639. URL: %3CGo%20to%20ISI%3E://WOS:000393878900013.

Skuk, V. G., R. Palermo, L. Broemer, and S. R. Schweinberger (2019). “Autistic traits are linked to individual differences in familiar voice identification”. In: *Journal of Autism and Developmental Disorders* 49.7, pp. 2747–2767. ISSN: 0162-3257. DOI: 10.1007/s10803-017-3039-y. URL: %3CGo%20to%20ISI%3E://WOS:000474414900010.