

Thema

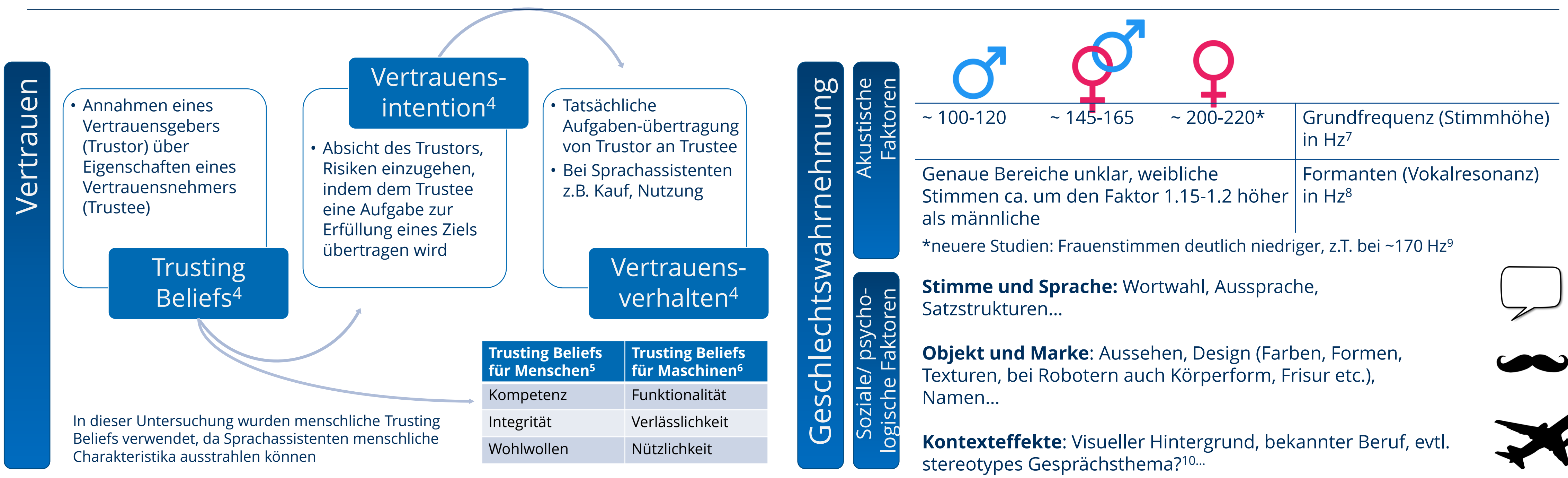
Theorie & Forschung

Methodik

Ergebnisse & Implikationen

Knapp 70% der aktuell existierenden Sprachassistentensysteme nutzen ausschließlich weibliche Stimmen. Assistenten mit männlichen Stimmen oder mit der Möglichkeit, zwischen verschiedenen Optionen zu wählen, sind dagegen deutlich in der Unterzahl¹. Technologien, die soziale Hinweisreize aussenden, werden gemäß des „Computers Are Social Actors“-Paradigmas wie soziale Akteure behandelt². Dies führt dazu, dass auf Sprachassistenten, die ein bestimmtes Geschlecht andeuten, beispielsweise Stereotype angewandt

werden können. Mit der Entwicklung der geschlechtsneutralen Stimme Q durch einen Zusammenschluss dänischer Firmen und Organisationen³ ergab sich 2019 die Frage, ob sich dadurch die Anwendung von Stereotypen umgehen lässt oder neutrale Stimmen im Gegenteil als unangenehm empfunden werden. Dies wurde in der durchgeführten Masterarbeit in Hinblick auf das Vertrauen in ein Sprachassistentensystem betrachtet.



Stimmen

Grundlagenstimme: de-DE WaveNet B (Google), Grundfrequenz 156 Hz, Formantenanalyse ergab Werte im Bereich zwischen bekannten männlichen und weiblichen Werten

Produktion von 3 männlichen, 3 weiblichen, 9 neutralen Stimme aus der Grundlagenstimme: Bearbeitung von Grundfrequenz (Festlegung der exakten Höhe) und Formanten (Veränderung der so erzeugten Stimme um verschiedene experimentelle Faktoren)

Pretest der Stimmen (Stichprobe: n = 33, Beurteilung von je 8 Stimmen)

Finale Stimmen	Grundfrequenz (Hz)	Formanten (Faktor)	Geschlechtswahrnehmung (1 = männlich, 7 = weiblich)
M	110	0.95	1.65
N	156	0.95	3.73
W	220	1	6.86

Untersuchung

Aufbau: Online-Experiment mit 9 Experimentalgruppen (3 Stimmen, 3 Themen (Flugzeuge, Pinguine, Liebe))

Ablauf:

- Erhebung verschiedener Kontrollvariablen
- Anhören der jeweiligen Stimme
- Erhebung abhängiger Variablen (empfundenes Vertrauen, wahrgenommenes Geschlecht, Einschätzung des Themas)
- Erhebung soziodemografischer Faktoren

Stichprobe: n = 343, 58% weiblich, Durchschnittsalter ca. 30 Jahre, grundlegend vertraut mit Sprachassistenten

Wahrnehmung der akustisch neutralen Stimme

Abb. 1: Direkter und indirekter Effekt des Themengeschlechts auf die Wahrnehmung der neutralen Stimme

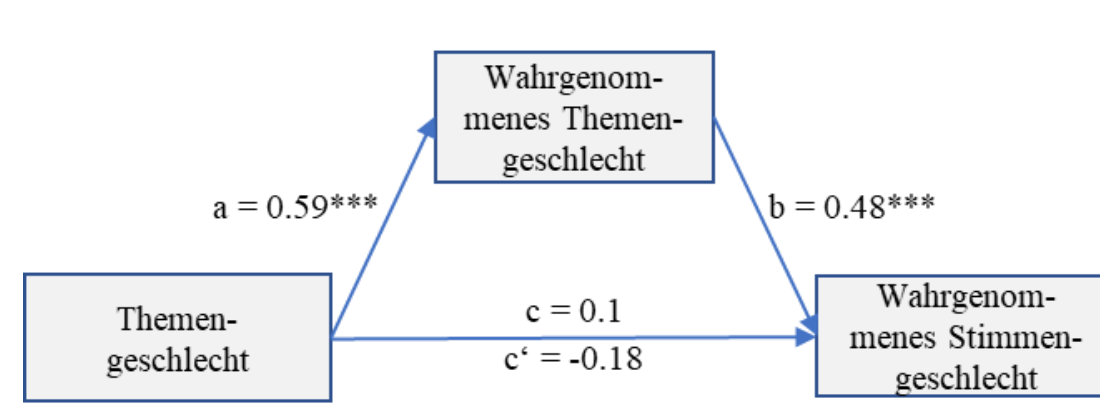


Abb. 2: Effekt des wahrgenommenen Themengeschlechts auf die Wahrnehmung der neutralen Stimme

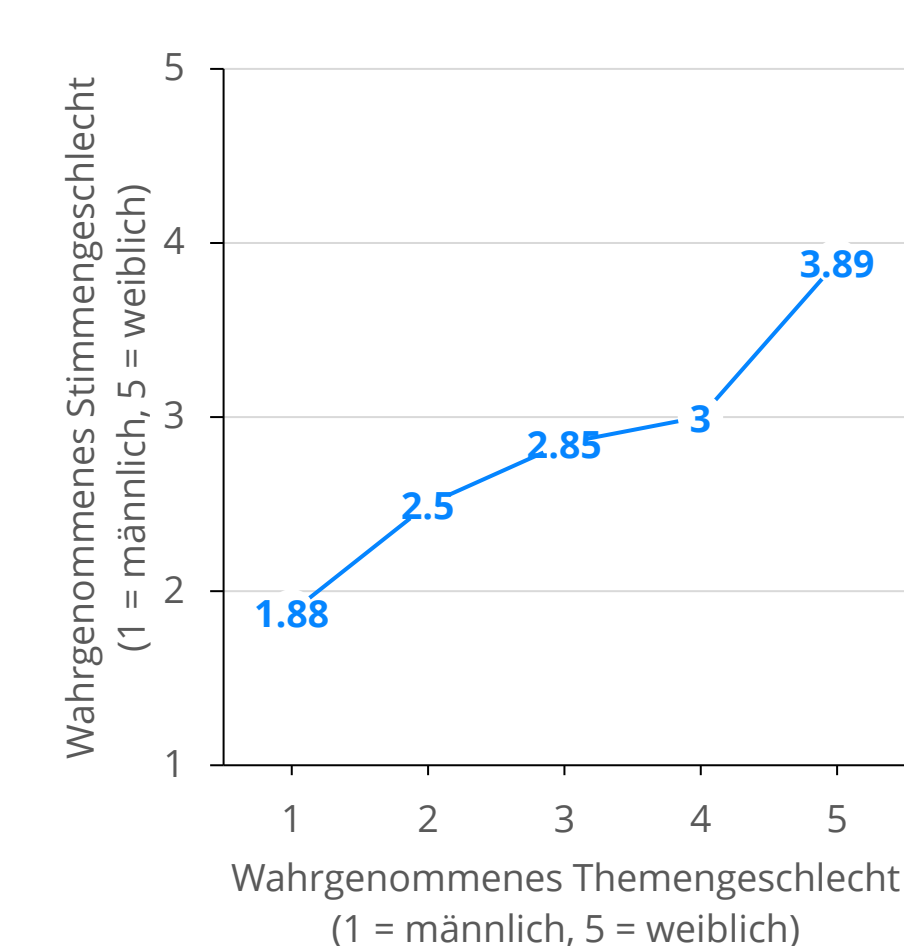
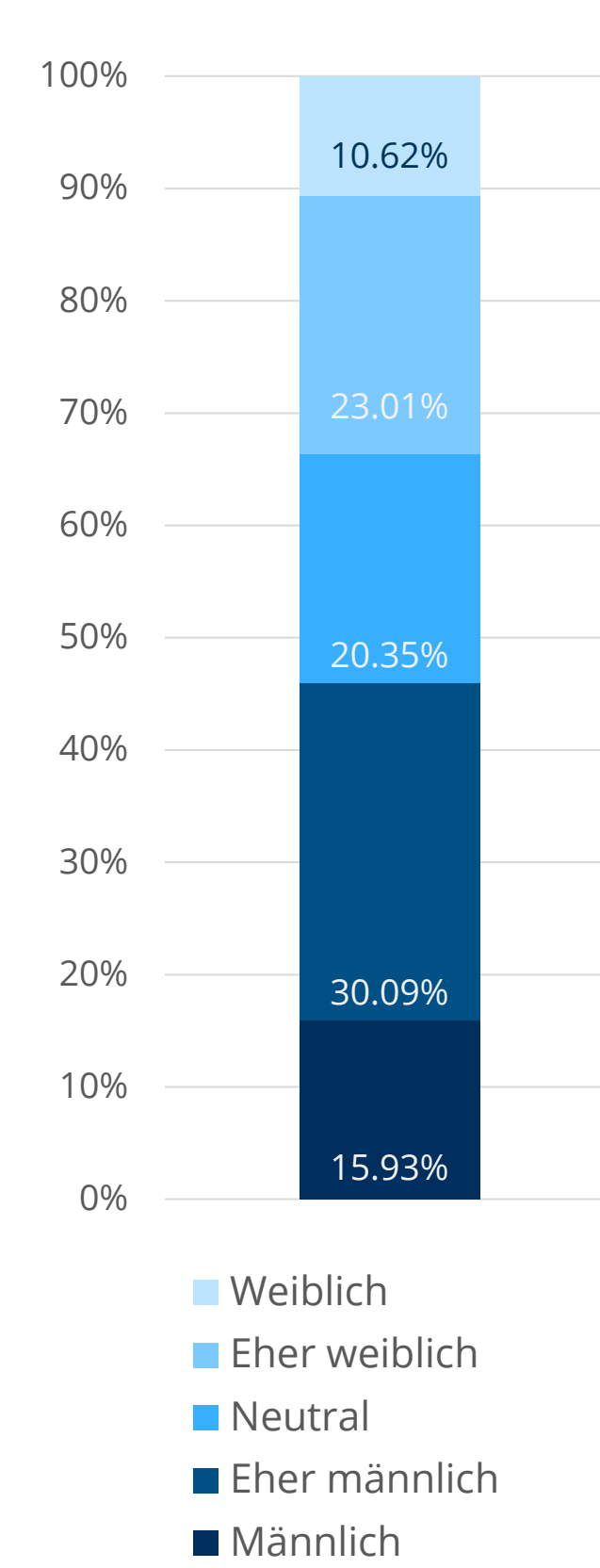
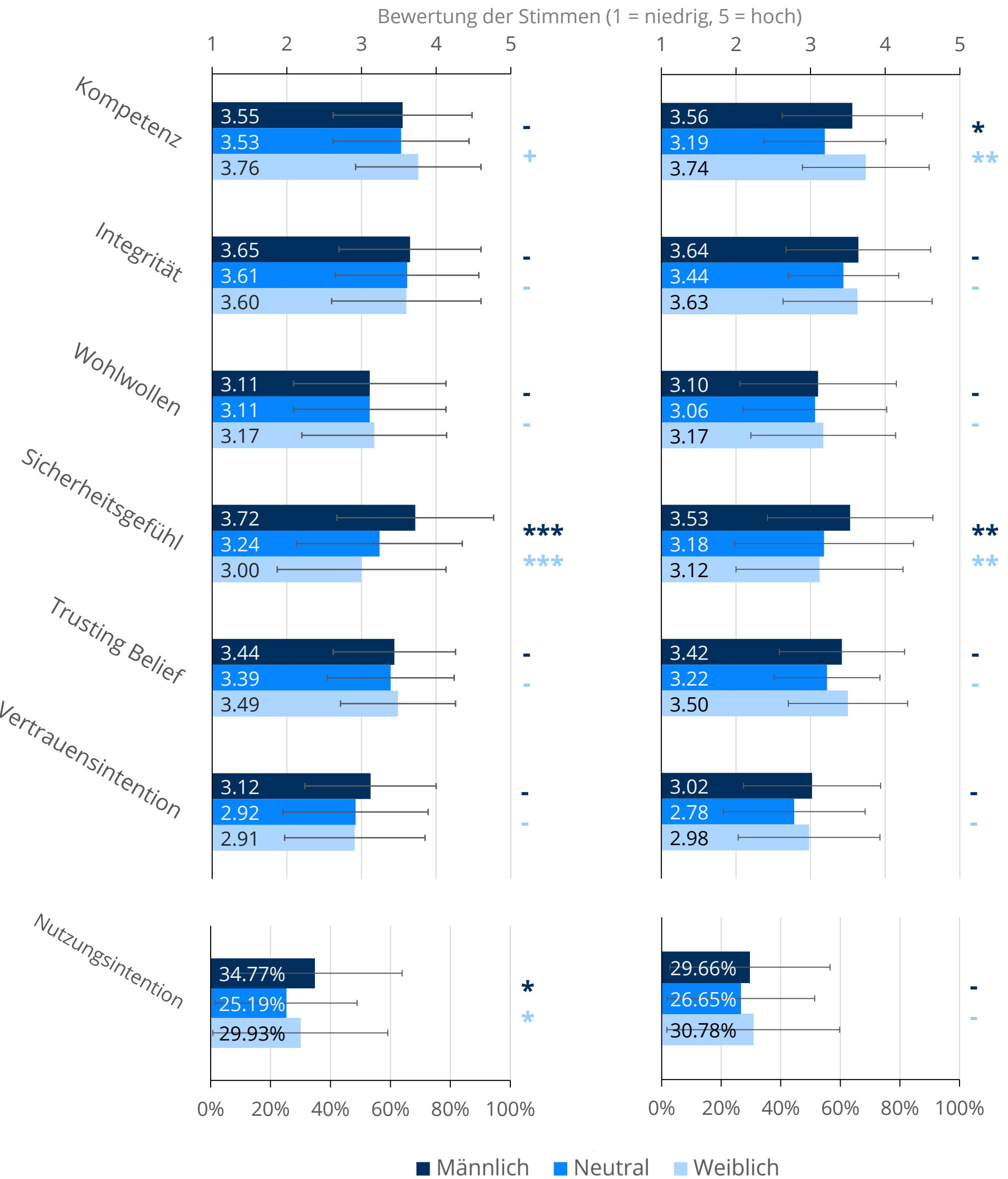


Abb. 3: Wahrnehmung der neutralen Stimme



Einfluss von Geschlecht auf Vertrauen

Abb. 4: Einfluss des akustischen Stimmengeschlechts (links) und des wahrgenommenen Stimmengeschlechts (rechts) auf das empfundene Vertrauen der Proband*innen



Beim Einfluss des Stimmengeschlechts auf die empfundene Vertrauenswürdigkeit zeigten sich nur begrenzte Effekte. Die neutrale Stimme wurde tendenziell negativer beurteilt als männliche und weibliche Stimmen; dies war zwar selten signifikant, aber bei fast jeder abhängigen Variable zu beobachten. Es zeigten sich auch Unterschiede zwischen dem akustischen und dem wahrgenommenen Stimmengeschlecht. Dieser Effekt ist vermutlich darauf zurückzuführen, dass die eher negative Wahrnehmung der akustisch neutralen Stimme

teilweise mit der positiveren Bewertung der akustisch männlichen und weiblichen Stimmen verrechnet wurde. Insgesamt deuten die Ergebnisse darauf hin, dass eine neutrale Stimme sich momentan nicht für den Einsatz in kommerziellen Sprachassistenten eignen würde: Insbesondere die Nutzungsintention ist bei einer akustisch neutralen Stimme signifikant niedriger als bei Stimmen, die eindeutig einem Geschlecht zugeordnet werden können.

1 Sey & Fesalbon (2019). "OK Google: Is AI Gendered?". In Sey & Hafkin (Hg.), Taking stock: Data and evidence on gender equality in digital access, skills and leadership (S. 144-145). United Nations University Institute on Computing and Society/International Telecommunications Union: Macau.

2 Reeves & Nass (1996). The media equation: How people treat computers, televisions, and new media like real people and places. CSLI Publ.

3 Genderless Voice (2020). genderlessvoice.com

4 McKnight, Cummings & Chervany (1998). Initial Trust Formation in New Organizational Relationships. Academy of Management Review, 23(3), 473-490.

5 Mayer, Davis & Schoorman (1995). An Integrative Model Of Organizational Trust. Academy of Management Review, 20(3), 709-734.

6 Lankton, McKnight & Tripp (2015). Technology, Humanness, and Trust: Rethinking Trust in Technology. Journal of the Association for Information Systems, 16(10), 880-918.

7 Simpson (2009). Phonetic differences between male and female speech. Language and Linguistics Compass, 3(2), 621-640. // Gelfer & Bennett (2013). Speaking fundamental frequency and vowel formant frequencies: Effects on perception of gender. Journal of Voice, 27(5), 556-566.

8 Hillenbrand, Getty, Clark & Wheeler (1995). Acoustic characteristics of American English vowels. The Journal of the Acoustical Society of America, 97(5 Pt 1), 3099-3111. // Peterson & Barney (1952). Control Methods Used in a Study of the Vowels. The Journal of the Acoustical Society of America, 24(2), 175-184. // Pätzold & Simpson (1997). Acoustic analysis of German vowels in the Kiel Corpus of Read Speech. Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung Universität Kiel(32), 215-247. // Wu & Childers (1991). Gender recognition from speech. Part I: Coarse analysis. The Journal of the Acoustical Society of America, 90(4 Pt 1), 1828-1840.

9 Berg, Fuchs, Wirkner, Loeffler, Engel & Berger (2017). The Speaking Voice in the General Population: Normative Data and Associations to Sociodemographic and Lifestyle Factors. Journal of Voice, 31(2), 257.e13-257.e24.

10 Sutton (2020). Gender Ambiguous, not Genderless. In M. I. Torres, S. Schlögl, L. Clark & M. Porcheron (Hg.), Proceedings of the 2nd Conference on Conversational User Interfaces (S. 1-8). ACM.