What happens when a virtual character speaks with the cloned voice of a person you know?

Johanna Magdalena Kuch, Daksitha Senel Withanage Don, Niklas Heimerl, Patrick Gebhard, Elisabeth André, Tanja Schneeberger

Angaben zur Veröffentlichung / Publication details:

Kuch, Johanna Magdalena, Daksitha Senel Withanage Don, Niklas Heimerl, Patrick Gebhard, Elisabeth André, and Tanja Schneeberger. 2026. "What happens when a virtual character speaks with the cloned voice of a person you know?" In *CHI EA '26: proceedings of the extended abstracts of the 2026 CHI Conference on Human Factors in Computing Systems, Barcelona, Spain, April 13-17, 2026*, edited by Nuria Oliver, David A. Shamma, Heloisa Candello, Pablo Cesar, Pedro Lopes, Valentino Artizzu, Fiona Draxler, et al., 737. New York, NY: ACM. https://doi.org/10.1145/3772363.3799332.

What Happens When a Virtual Character Speaks With the Cloned Voice of a Person You Know?

Johanna Magdalena Kuch
Human-Centered Artificial Intelligence
Augsburg University
Augsburg, Germany
johanna.kuch@uni-a.de

Daksitha Senel Withanage Don
Chair for Human-Centered Artificial Intelligence
Augsburg University
Augsburg, Germany
daksitha.withanage.don@uni-a.de

Niklas Heimerl
Human-Centered Artificial Intelligence
Augsburg University
Augsburg, Germany
niklas.heimerl@uni-a.de

Patrick Gebhard
Intelligent User Interfaces
German Research Center for Artificial Intelligence
Saarbrücken, Saarland, Germany
patrick.gebhard@dfki.de

Elisabeth André
Human-Centered Artificial Intelligence
Augsburg University
Augsburg, Germany
andre@informatik.uni-augsburg.de

Tanja Schneeberger
German Research Center for Artificial Intelligence
Saarbrücken, Saarland, Germany
tanja.schneeberger@dfki.de

Abstract

What happens when a virtual character speaks with the cloned voice of a person you know? Our results suggest that such a known-voice clone can subtly shape social perception, but does not reliably transfer the perceived personality of the original speaker to the virtual character. We investigate whether voice cloning can shape the attribution of personality and social traits to a virtual character when that character speaks with a cloned version of a known voice. Across two studies, we compared ratings of an original voice speaker with ratings of a virtual character speaking with either a known-voice clone or an unknown voice. The first study was a pilot in which students were exposed to the voice of their lecturer. The second study was a preregistered follow-up in which content creators' voices were presented to their respective communities. Participants evaluated personality traits and social characteristics, and we measured how similar these ratings were to those of the original speaker. Quantitative results revealed small and selective effects. In the pilot study, differences emerged for likability and agreeableness, while in the follow-up, effects were observed for extraversion. Qualitative analyses showed a shift in focus away from perceptions of the salience of artificiality and unnaturalness toward the incongruence between voice and visual appearance. Overall, the findings indicate that *known-voice cloning* is not a general mechanism for personality transfer. Instead, it functions as a weak and context-dependent cue whose influence is constrained by multimodal coherence and users' expectations. These results provide practical guidance on when *known-voice cloning* may be beneficial, when it may backfire, and which aspects should be prioritized in the design of embodied voice interfaces.

CCS Concepts

• **Human-centered computing** → **User studies**; Sound-based input / output; **Empirical studies in HCI**; **Empirical studies in interaction design**.

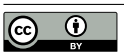
Keywords

voice cloning, known voice, personality attribution, social perception, embodied virtual characters, trust in artificial agents, uncanny valley, audiovisual congruence

ACM Reference Format:
Johanna Magdalena Kuch, Daksitha Senel Withanage Don, Niklas Heimerl, Patrick Gebhard, Elisabeth André, and Tanja Schneeberger. 2026. What Happens When a Virtual Character Speaks With the Cloned Voice of a Person You Know?. In *Extended Abstracts of the 2026 CHI Conference on Human Factors in Computing Systems (CHI EA '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3772363.3799332

1 Introduction

Voice is a strong cue in social perception and shapes impressions of virtual characters [5, 6, 8, 13]. Voice cloning makes it easy to design voices by reusing voices from real people. With zero-shot voice cloning, realistic speech can be generated from short reference recordings, allowing virtual characters to speak with clones of *known voices*. What remains largely unexplored is how this design choice affects social perception. Prior work on voice cloning has largely focused on cloning a user's own voice [9, 11]. In contrast, much less is known about what happens when a virtual character speaks with the cloned voice of someone users already know. Because people hold rich, person-specific representations of familiar individuals, including personality and social attributes [1, 28], a known voice may subtly shape how a virtual character is perceived. This raises the question of whether such person-based attributions extend to an unfamiliar character when it speaks with a *known-voice clone*. This question is especially relevant today, since people are repeatedly exposed to the voices of specific public figures through social media platforms. If a *known voice* biases social perception of virtual characters, voice cloning could become a subtle mechanism for shaping impressions in highly personalized digital spaces. We



CHI EA '26, Barcelona, Spain

ACM ISBN 979-8-4007-2281-3/26/04
https://doi.org/10.1145/3772363.3799332

investigate these questions through two studies in distinct exposure contexts: a pilot study involving student ratings of a university lecturer and a preregistered follow-up involving ratings of content creators by their respective audiences. We address the following research question:

RQ: To what extent does using a cloned voice of a familiar person influence trait and social evaluations of a virtual character relative to an unfamiliar voice, and how consistent are such influences across traits and contexts?

In this work, we conceptualize attribution transfer as the degree of alignment between memory-based trait evaluations of a known person and evaluations of a virtual character speaking with a cloned version of that person's voice.

This paper makes the following contributions: We provide empirical evidence across two studies that using a cloned voice of a familiar person does not produce robust or generalizable alignment between evaluations of a virtual character and memory based evaluations of the original speaker. Instead using a known-voice clone can yield small and selective effects on specific traits, but that these effects are inconsistent across different contexts. We show that *known-voice cloning* primarily shifts users' perceptual focus toward speaker identity and voice–appearance congruence, while issues of visual rigidity and vocal artificiality remain dominant. We report results from a preregistered follow up study, strengthening the robustness and transparency of the findings.

2 Related Work

Voices as Cues for Person-Based Social Attributions. From short voice samples, listeners form social judgments, including personality impressions and evaluations such as trustworthiness and competence [15, 16, 19].

When a voice is a *known voice*, hearing it can trigger person-specific representations or associated attributions even without explicit identification [2, 14]. This suggests that social perception can be biased by retrieved person information beyond acoustic cues alone.

Transfer of Person-Based Characteristics to Unknown Characters. Observers can project person-based attributions from familiar individuals onto new targets when a cue activates an existing representation [10, 27]. Voice-based person knowledge can bias judgments even without explicit speaker identification, suggesting a pathway for attribution transfer through *known voices* [2, 11].

Voice Cloning as an Enabling Technology. Zero-shot multi-speaker text-to-speech can imitate previously unseen voices from short reference recordings [4, 22]. This enables deploying known-voice clones at scale, but it remains unclear whether they transfer person-based attributions to embodied virtual characters, and whether this differs between in-person and media-based contexts.

3 Overview of Studies

In both studies, participants first provided memory ratings of a known person and then rated the same virtual character in two voice conditions, *known-voice clone* and *unknown voice*. Attribution transfer was operationalized as the absolute distance between person ratings and virtual character ratings, where smaller distances indicate stronger transfer. Based on prior work on person-based activation through familiar cues, we formulated a directional hypothesis:

H1: A virtual character speaking with a known-voice clone will show reduced distance to memory-based ratings of the known person compared to the same character speaking with an unknown voice.

Formally, we tested the directional hypothesis that
$H_1 : D_{\text{known-voice clone}} < D_{\text{unknown voice}}$,
where $D = |R_{\text{person}} - R_{\text{character}}|$.

Study 1 was a pilot study in a university context (lecturer). **Study 2** was preregistered and tested the same distance-based hypothesis in an online media context (content creators) with a larger sample.

Design. Both studies followed a within subject design with three conditions. The lecturer or content creator condition was always presented first to capture participants' baseline memory based impressions before any potential influence from exposure to the virtual character. This was followed by the two virtual character conditions in random order:(1) *Lecturer/Content Creator*: ratings of the real person from memory. (2) *Known-voice clone*: virtual character speaking with a voice clone of the lecturer/content creator. (3) *Unknown voice*: the same character speaking with a cloned voice of an unknown person. The spoken content was identical across voice conditions. Appearance and content were held constant while only voice differed; the virtual character and lip synchronization were created using MeLaX [30]. Age, gender, and ethnicity were aligned with the respective content creators to reduce visual mismatch that could obscure attribution transfer effects [6].

4 Pilot Study with Lecturer

Participants. We recruited students from the lecturer's course. Participation was voluntary and took place during lecture and exercise sessions on participants' own devices while the lecturer was not present. Eighteen complete datasets were collected; one participant was excluded due to insufficient familiarity with the lecturer, resulting in a final sample of $n = 17$ (9 men, 5 women, 1 non-binary, 2 did not report gender; age: $M = 22.46$, $SD = 3.02$, range = 20–32). Participants were predominantly enrolled in computer science or related fields.

Stimuli. We presented short video clips (approximately 15 s) of a virtual character speaking phonetically balanced sentences [12]. Voice clones were generated using *OpenVoice* [22]. An example frames of the virtual character is provided in the Appendix.

Measures. We measured Big Five other-ratings using the German Big Five Inventory [23] adapted to rate others [29] (21 items), as well as likability (5 items) [7] and competence (3 items) [3]. Big Five items used 5-point Likert scales, and likability and competence items used 7-point Likert scales. Negatively keyed items were reverse coded. Full item wordings are provided in the Appendix. Open-ended questions captured participants' qualitative impressions of the lecturer and the virtual character. Responses were analyzed using qualitative content analysis following Mayring [18]. Question wordings are provided in the Appendix.

Results. We tested the directional hypothesis with one-tailed Wilcoxon signed-rank tests. Distances were smaller for likability ($W = 11.5$, $p = .01$, $r = .75$) and agreeableness ($W = 16.0$, $p =$

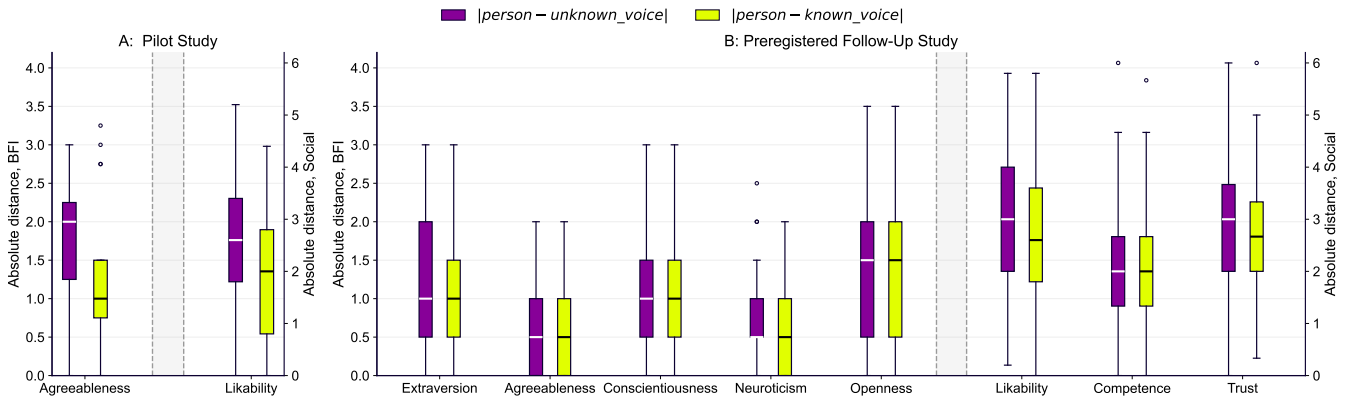


Figure 1: Absolute distance scores between memory based person ratings and virtual character ratings by voice condition across both studies. Panel A shows the pilot study with agreeableness on the left y axis and likability on the right y axis. Panel B shows the follow up study with Big Five traits on the left y axis and social evaluation measures on the right y axis.

.038, $r = .59$) in the *known-voice clone* condition (Figure 1). No significant differences emerged for competence or the remaining Big Five dimensions. Effect sizes for all comparisons, including non-significant results, are reported in the Appendix. Free-text responses in the pilot study distinguished clearly between the real lecturer and the two virtual character conditions. The lecturer was predominantly described as competent, knowledgeable, and socially positive, whereas both virtual characters caused more ambivalent impressions. The *known-voice clone* was perceived as competent but emotionally restrained and socially mixed. The *unknown voice condition* showed the more negative results in social and emotional descriptions. Across both virtual character conditions, participants frequently reported difficulty making personality judgments based on very short video clips and deficiencies in sound and animation quality. Selective effects in the pilot motivated a follow-up study with longer stimuli in an online context and the addition of trust [25] as an outcome.

5 Follow-Up Study with Content Creators

5.1 Methods

Preregistration and Design. The study was preregistered on AsPredicted [1]. Within-subject design was used as previously described. Two content creators were included, and participants evaluated the creator they reported greater prior exposure to.

Participants. Participants were recruited online (Reddit, Instagram, Newsletters). Responses were excluded if they were incomplete, failed attention checks, or indicated insufficient prior exposure to the selected content creator, resulting in 65 datasets retained for analysis. The final sample of the follow-up study consisted of $n = 65$ participants. Participants ranged in age from 19 to 52 years ($M = 27.8$, $SD = 6.3$). Gender was self-reported by 30 men and 32 women, while three participants did not report their gender.

Stimuli. The virtual character was refined after the pilot study with adjustments to camera angle and visual appearance. Pilot feedback suggested that very short, content-poor sentences did not support differentiated judgments. We therefore used longer, neutral statements to provide minimal context for social perception ratings [21]. Voice clones were generated from short reference recordings. We chose a different cloning system than in the pilot to achieve better clone quality [4]

Measures. Outcomes were Big Five other-ratings assessed with brief scales [24], social evaluation measures for likability, competence, and trust [25], and free text answers, collected identically for creator ratings and each virtual character voice block. Prior exposure to the voice was required via screening but not directly measured as a continuous predictor. All participants provided informed consent; data were collected and stored in anonymized form in accordance with applicable data protection requirements.

5.2 Results

A total of $n = 208$ participants initially started the follow-up study. Eighty-eight participants exited the study prematurely. Of the remaining participants, 15 indicated that they were not familiar with the content creator and were excluded. An additional 14 participants failed a knowledge check verifying familiarity with the content creator and were excluded. Of the remaining participants, 69 passed two embedded instruction checks requiring a specific response format. Three participants reported no consumption of the content creator within the past three months and were excluded. Finally, one participant was excluded due to non-interpretable open-ended responses. The final sample consisted of $n = 65$ participants.

Preregistered tests. The preregistered tests directly examined H1 across all specified outcomes. We used paired tests appropriate to the distribution of the difference scores and report effect sizes accordingly. Only extraversion showed a significant effect in the expected direction ($p = .041$, $r = .39$, $n = 65$); all other outcomes were not significant (Figure 1).

Qualitative findings. Consistent with the pilot study, participants frequently commented on visual rigidity and unnatural facial movement of the virtual character. In the known-voice clone condition, responses were dominated by explicit recognition of voice similarity to the content creator, whereas this theme was absent

[1] https://aspredicted.org/zy8j9m.pdf

in the unknown voice condition. Voice artificiality and monotony were commonly reported in both conditions, while descriptions of the voice as comparatively natural or expressive occurred slightly more often for the unknown voice. Mentions of incongruence between voice and appearance were present in both conditions, with a modest emphasis in the known-voice clone condition. Some participants reported that voice familiarity slightly attenuated negative impressions without reversing them. Overall, the qualitative patterns suggest that *known-voice cloning* shifts perceptual focus toward speaker identity and voice–appearance fit, while visual rigidity and vocal artificiality remain salient across conditions.

6 General Discussion, Limitations and Future Work

Across two studies, we tested whether a *known-voice clone* enables attribution transfer to a virtual character, operationalized as reduced absolute distances between memory-based person ratings and character ratings. The evidence points to small, selective effects rather than robust transfer. Reduced distance scores should be interpreted cautiously, as they may also reflect general judgment uncertainty or limited differentiation under brief exposure rather than genuine person-specific attribution transfer. In **Study 1**, distances were smaller for likability and agreeableness in the *known-voice clone* condition, whereas no effects emerged for other traits. In **Study 2**, extraversion showed the hypothesized effect, while social evaluation measures did not replicate the pilot findings. Qualitative responses suggest that the *known voice* is often noticed, but does not consistently translate into more person-consistent evaluations. Instead, participants frequently highlighted visual rigidity and vocal artificiality across conditions, and comments in the *known-voice clone* condition often centered on speaker identity and voice-appearance congruence. The isolated extraversion effect is consistent with prior work indicating that extraversion is comparatively voice-diagnostic and can be inferred from prosodic cues [20, 26]. In contrast, other traits and social evaluations may require richer behavioral information or longer exposure [17, 31]. Our findings suggest a constrained transfer account: a familiar cloned voice may increase identity salience, but perceived artificiality and audiovisual mismatch limit consistent trait alignment.

Several limitations should be considered when interpreting the findings. Across both studies, salient perceptual characteristics of the virtual character, including visual rigidity and vocal artificiality, likely constrained attribution transfer. As discussed above, these cues often dominated participants' evaluations and may have reduced the relative impact of using the known-voice clone. The two study contexts differ in how the target persons are typically experienced. Lecturers are encountered in physical settings, whereas content creators are primarily experienced through screen-based media. This difference may influence the baseline similarity between remembered targets and the virtual character, and thus the sensitivity of the distance-based transfer measure across contexts. Participants were screened for prior exposure to the target voice, but familiarity and recognition strength were not measured in detail. Judgments were based on brief video clips and, in Study 2, on short personality scales, which may have limited the formation and detection of stable attribution patterns. In addition, the sample size may have limited the power to detect small effects. We did not include a dedicated manipulation check of perceived voice naturalness. Future work should systematically assess perceived naturalness and audiovisual congruence to disentangle familiarity effects from general realism effects.

This work suggests three clear directions for future research. Rather than treating voice cloning as an isolated feature, future work should investigate it as part of an integrated design process for embodied characters. Our findings motivate the **development and evaluation of a design-oriented tool** that uses voice cloning to systematically explore and evaluate voice choices in relation to visual appearance, expressiveness, and overall character coherence. Such a tool could support designers in comparing cloned and non-cloned voices early in the design process, identifying potential mismatches, and making informed decisions about when voice familiarity is likely to support or undermine social perception. Future studies should allow longer or interactive exposures to virtual characters. Such settings may be necessary for stable person-based attributions to emerge and for voice-based effects to unfold beyond initial impressions. Future work should directly measure voice familiarity and recognition strength, and model individual differences in prior knowledge and expectations. This would help clarify whether known-voice effects are driven by a subset of users with strong person representations and under which conditions these effects generalize. Future work should also evaluate alternative transfer measures beyond absolute distance scores.

7 Conclusion

Across two studies, we tested whether using a known-voice clone based on prior voice exposure enables attribution transfer from a known person to a virtual character. The results indicate small, outcome-specific effects rather than reliable personality transfer across traits and contexts. While isolated effects emerged for likability, agreeableness, and extraversion, these did not generalize consistently. Taken together, *known-voice cloning* should be understood as a subtle and context-dependent design cue rather than a mechanism for robust attribution transfer. For the design of embodied voice interfaces, this suggests that cloning a known voice is unlikely to be beneficial in isolation. Instead, voice cloning is best suited as a diagnostic and exploratory instrument for anticipating perceptual consequences of voice choices within a multimodal character design space.

Acknowledgments

We thank the content creators and all participants for their time and feedback. We used AI-based tools to assist with language editing and the preparation and formatting of tables. These tools were not used to generate study ideas, analyses, or substantive interpretations. All scientific decisions, data analyses, and final manuscript wording were reviewed, revised, and approved by the authors.

References

[1] Galen V. Bodenhausen and Andrew R. Todd. 2010. Social cognition. *Wiley interdisciplinary reviews. Cognitive science* 1, 2 (2010), 160–171. doi:10.1002/wcs.28

[2] Sandra Campeanu, Fergus I. M. Craik, and Claude Alain. 2015. Speaker's voice as a memory cue. *International journal of psychophysiology : official journal of the*

International Organization of Psychophysiology 95, 2 (2015), 167–174. doi:10.1016/j.ijpsycho.2014.08.988

[3] Linda L. Carli, Suzanne J. LaFleur, and Christopher C. Loeber. 1995. Nonverbal behavior, gender, and influence. *Journal of Personality and Social Psychology* 68, 6 (1995), 1030–1041. doi:10.1037/0022-3514.68.6.1030

[4] Edresson Casanova, Kelly Davis, Eren Gölge, Görkem Göknar, Iulian Gulea, Logan Hart, Aya Aljafari, Joshua Meyer, Reuben Morais, Samuel Olayemi, and Julian Weber. 2024. XTTS: a Massively Multilingual Zero-Shot Text-to-Speech Model. In *Interspeech 2024*. ISCA, ISCA, 4978–4982. doi:10.21437/Interspeech.2024-2016

[5] Erin K. Chiou, Noah L. Schroeder, and Scotty D. Craig. 2020. How we trust, perceive, and learn from virtual humans: The influence of voice quality. *Computers & Education* 146 (2020), 103756. doi:10.1016/j.compedu.2019.103756

[6] Minsoo Choi, Alexandros Koilias, Matias Volonte, Dominic Kao, and Christos Mousas. [n. d.]. Exploring the Appearance and Voice Mismatch of Virtual Characters. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 555–560. doi:10.1109/ISMAR-Adjunct60411.2023.00118

[7] Joanna Hale and Antonia F. de C. Hamilton. 2016. Cognitive mechanisms for responding to mimicry from others. *Neuroscience and biobehavioral reviews* 63 (2016), 106–123. doi:10.1016/j.neubiorev.2016.02.006

[8] Darragh Higgins, Katja Zibrek, Joao Cabral, Donal Egan, and Rachel McDonnell. 2022. Sympathy for the digital: Influence of synthetic voice on affinity, social presence and empathy for photorealistic virtual humans. *Computers & Graphics* 104 (2022), 116–128. doi:10.1016/j.cag.2022.03.009

[9] Oliver Jaggy, Stephan Schwan, and Hauke S. Meyerhoff. 2025. AI-determined similarity increases likability and trustworthiness of human voices. *PloS one* 20, 3 (2025), e0318890. doi:10.1371/journal.pone.0318890

[10] Michael W. Kraus and Serena Chen. 2010. Facial-feature resemblance elicits the transference effect. *Psychological science* 21, 4 (2010), 518–522. doi:10.1177/0956797610364949

[11] Johanna Magdalena Kuch, Marcel Heisler, Stina Klein, Silvan Mertes, Lennart Eing, Elisabeth André, and Christian Becker-Asano. 2025. Your Robot, My Voice: Enhancing Android Robot Likability through Personalization by Cloning the User's Voice. In *2025 34th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 192–198. doi:10.1109/RO-MAN63969.2025.11217611

[12] Volker Kuehnel, Birger Kollmeier, and Kirsten Wagener. 1999. Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests. *Zeitschrift für Audiologie* 38 (1999), 4–15.

[13] Luchcha Lam, Minsoo Choi, Magzhan Mukanova, Klay Hauser, Fangzheng Zhao, Richard Mayer, Christos Mousas, and Nicoletta Adamo-Villani. 2023. Effects of Body Type and Voice Pitch on Perceived Audio-Visual Correspondence and Believability of Virtual Characters. In *ACM Symposium on Applied Perception 2023*, Alex Chapiro, Andrew Robb, Funda Durupinar, Qi Sun, and Lauren Buck (Eds.). ACM, New York, NY, USA, 1–11. doi:10.1145/3605495.3605791

[14] Nadine Lavan. 2025. Time after time: Voice perception from first impressions to identity recognition. *Quarterly Journal of Experimental Psychology* 78, 12 (2025), 2583–2593. doi:10.1177/17470218251379036

[15] Nadine Lavan, Paula Rinke, and Mathias Scharinger. 2024. The time course of person perception from voices in the brain. *Proceedings of the National Academy of Sciences of the United States of America* 121, 26 (2024), e2318361121. doi:10.1073/pnas.2318361121

[16] Nadine Lavan and Clare A. M. Sutherland. 2024. Idiosyncratic and shared contributions shape impressions from voices and faces. *Cognition* 251 (2024), 105881. doi:10.1016/j.cognition.2024.105881

[17] Jeffrey A. Lepine, Jason A. Colquitt, and Amir Eerez. 2000. ADAPTABILITY TO CHANGING TASK CONTEXTS: EFFECTS OF GENERAL COGNITIVE ABILITY, CONSCIENTIOUSNESS, AND OPENNESS TO EXPERIENCE. *Personnel Psychology* 53, 3 (2000), 563–593. doi:10.1111/j.1744-6570.2000.tb00214.x

[18] Philipp Mayring. 2015. Qualitative Content Analysis: Theoretical Background and Procedures. In *Approaches to qualitative research in mathematics education*, Angelika Bikner-Ahsbahs, Christine Knipping, and Norma Presmeg (Eds.). Springer, Dordrecht and Heidelberg and New York and London, 365–380. doi:10.1007/978-94-017-9181-6_13

[19] Gelareh Mohammadi and Alessandro Vinciarelli. 2012. Automatic Personality Perception: Prediction of Trait Attribution Based on Prosodic Features. *IEEE Transactions on Affective Computing* 3, 3 (2012), 273–284. doi:10.1109/T-AFFC.2012.5

[20] Gelareh Mohammadi, Alessandro Vinciarelli, and Marcello Mortillaro. 2010. The voice of personality. In *Proceedings of the 2nd international workshop on Social signal processing*, A. Vinciarelli, Maja Pantic, and Alex Pentland (Eds.). ACM, New York, NY, USA, 17–20. doi:10.1145/1878116.1878123

[21] Lena Nadarevic and Edgar Erdfelder. 2017. Initial judgment task and delay of the final validity-rating task moderate the truth effect. doi:10.17605/OSF.IO/EUT35

[22] Zengyi Qin, Wenliang Zhao, Xumin Yu, and Xin Sun. 03.1. OpenVoice: Versatile Instant Voice Cloning. https://arxiv.org/pdf/2312.01479

[23] Beatrice Rammstedt and Oliver P. John. 2005. Kurzversion des Big Five Inventory (BFI-K). *Diagnostica* 51, 4 (2005), 195–206. doi:10.1026/0012-1924.51.4.195

[24] B. Rammstedt, C. J. Kemper, M. C. Klein, C. Beierlein, and A. Kovaleva. 2014. Big Five Inventory (BFI-10). doi:10.6102/ZIS76

[25] Daniel Roth, David Mal, Christian Felix Purps, Peter Kullmann, and Marc Erich Latoschik. 2018. Injecting Nonverbal Mimicry with Hybrid Avatar-Agent Technologies. In *Proceedings of the Symposium on Spatial User Interaction (ACM Conferences)*. ACM, New York, NY, 69–73. doi:10.1145/3267782.3267791

[26] Klaus R. Scherer. 1978. Personality inference from voice quality: The loud voice of extroversion. *European Journal of Social Psychology* 8, 4 (1978), 467–487. doi:10.1002/ejsp.2420080405

[27] Nurit Tal-Or, Shani Sela, Israel Igumnov, Hanoch Dov Milwidsky, Benjamin Rafaeli, and Michael Sanilevich. 2021. Does What We Know About Actors' Real Lives Influence Our Reactions to the Characters They Play? *Journal of Media Psychology* 33, 3 (2021), 155–164. doi:10.1027/1864-1105/a000293

[28] Mark A. Thornton and Jason P. Mitchell. 2017. Consistent Neural Activity Patterns Represent Personally Familiar People. *Journal of cognitive neuroscience* 29, 9 (2017), 1583–1594. doi:10.1162/jocn_a_01151

[29] Yannik Wiechers and Christian Kandler. 2025. Validierung des Deutschen Big Five Inventar-2. *Diagnostica* 71, 2 (2025), 64–75. doi:10.1026/0012-1924/a000344

[30] Daksitha Senel Withanage Don, Thomas Kiderle, Silvan Mertes, Dominik Schiller, Hannes Ritschel, and Elisabeth André. 2025. MeLaX: Conversations with Generative AI in Socially Interactive Agents. In *Companion Proceedings of the 30th International Conference on Intelligent User Interfaces (ACM Digital Library)*, Toby Li, Fabio Paternò, Kaisa Väänänen, Luis Leiva, Davide Spano, and Katrien Verbert (Eds.). Association for Computing Machinery, Erscheinungsort nicht ermittelbar, 163–166. doi:10.1145/3708557.3716363

[31] Qiaoning Zhang, X. Jessie Yang, and Lionel P. Robert. 2025. Voice Similarity and its Impact on Cognitive and Affective Trust in Automated Vehicles. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 69, 1 (2025), 1040–1045. doi:10.1177/10711813251364804