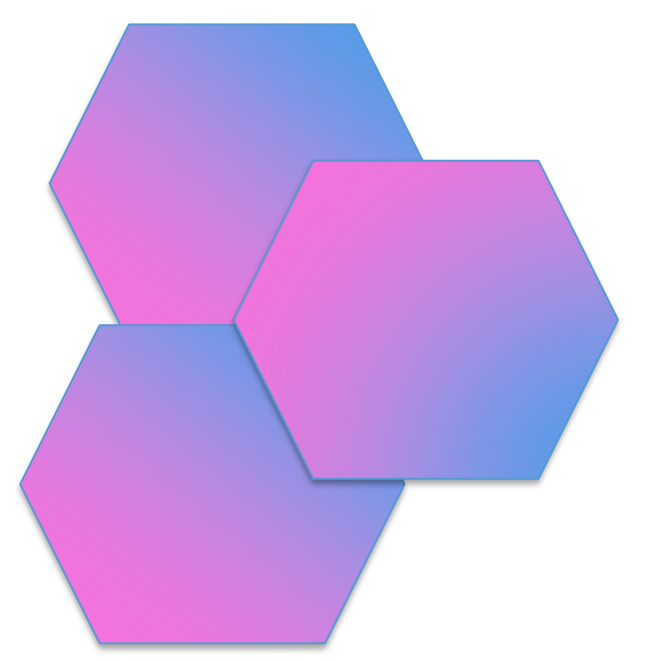


Do we blame it for its gender?

How specific gender cues affect the evaluation of virtual online assistants



Carolin Straßmann, Annika Arndt, Anna Dahm, Dennis Nissen, Björn Zwickler, Bejko Regy, Melissa Güven, Simon Schulz, Sabrina Eimler



Relevance and Research Question

Virtual online assistants give us recommendations on websites or help us with our daily lives. These agents have mostly a humanoid design and are associated with a gender. Based on the Media Equation theory (Reeves & Nass, 1996) assistants trigger the same social responses as humans. Consequently, gender stereotypes are applied, which were found to affect the perception of the agent (c.f. Nowak & Fox, 2018). The gender of the assistant can be conveyed by different cues, which might make these stereotypes more or less salient. The present study aims to investigate the effect of different cues representing assistant's gender on its evaluation.

Method

An online experiment with a 2x3x2 between-subjects-design, where gender (male vs. female), gender cues (either represented by name, embodied character or voice) and interaction quality (flawless interactions or incorrect interaction) was conducted. Participants completed the questionnaire and evaluated the assistant afterwards with regard to warmth and competence (Fiske, Cuddy, Glick & Xu, 2002)

44 female
112 male
3 diverse
N = 138
 $M_{age} = 23.93$
SD = 8.19



Illustration 1: Stimulus material

Hypotheses

- H1:** Male assistance systems are evaluated more competent.
- H2:** If a command is not recognized correctly, female assistance systems are evaluated less competent as male assistance systems.
- H3:** How do different gender cues effect the perceived warmth and competence of virtual online assistants systems.
- H4:** Women are perceived warmer than men.

Results

Results indicate that female assistants were perceived as warmer than male ($F(1, 135) = 4.58, p = .034, \eta^2 = 0.03$) and that when the interaction was flawless, the agent got evaluated as more competent after a flawless interaction than after an incorrect interaction ($F(1, 135) = 4.07, p = .046, \eta^2 = 0.03$). Moreover, the representation of the gender differed with regard to warmth ($F(2, 135) = 4.76, p = .010, \eta^2 = 0.07$), where the voice was perceived as significantly less warm than the name or embodied character. Additionally, a 3-way interaction between all independent variables occurred with regard to warmth ($F(2, 135) = 3.20, p = .044, \eta^2 = 0.05$): For female agents represented through an embodied character an interaction with a failure leads to a higher warmth evaluation than a flawless interaction.

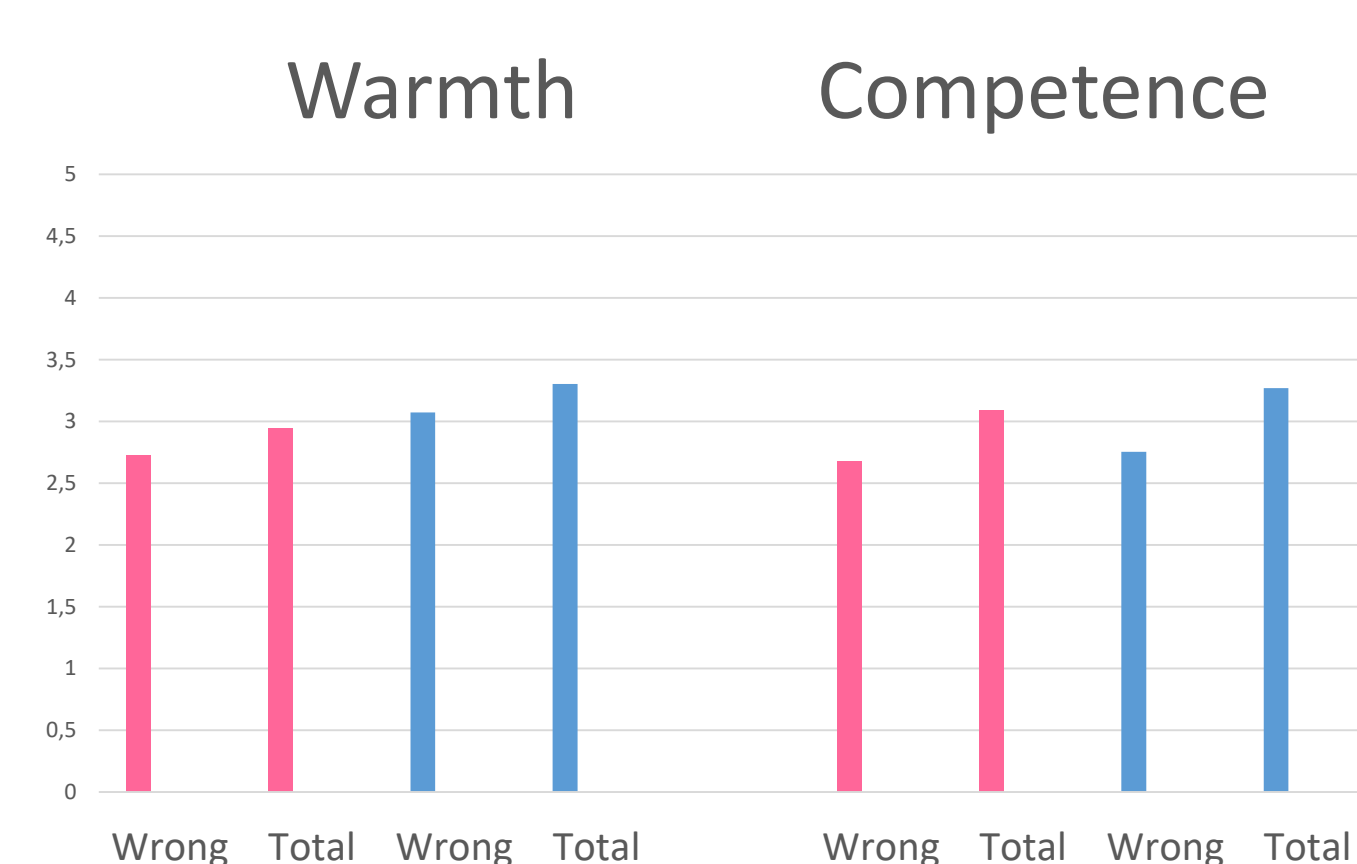


Illustration 2: Evaluation of assistance systems towards warmth / competence

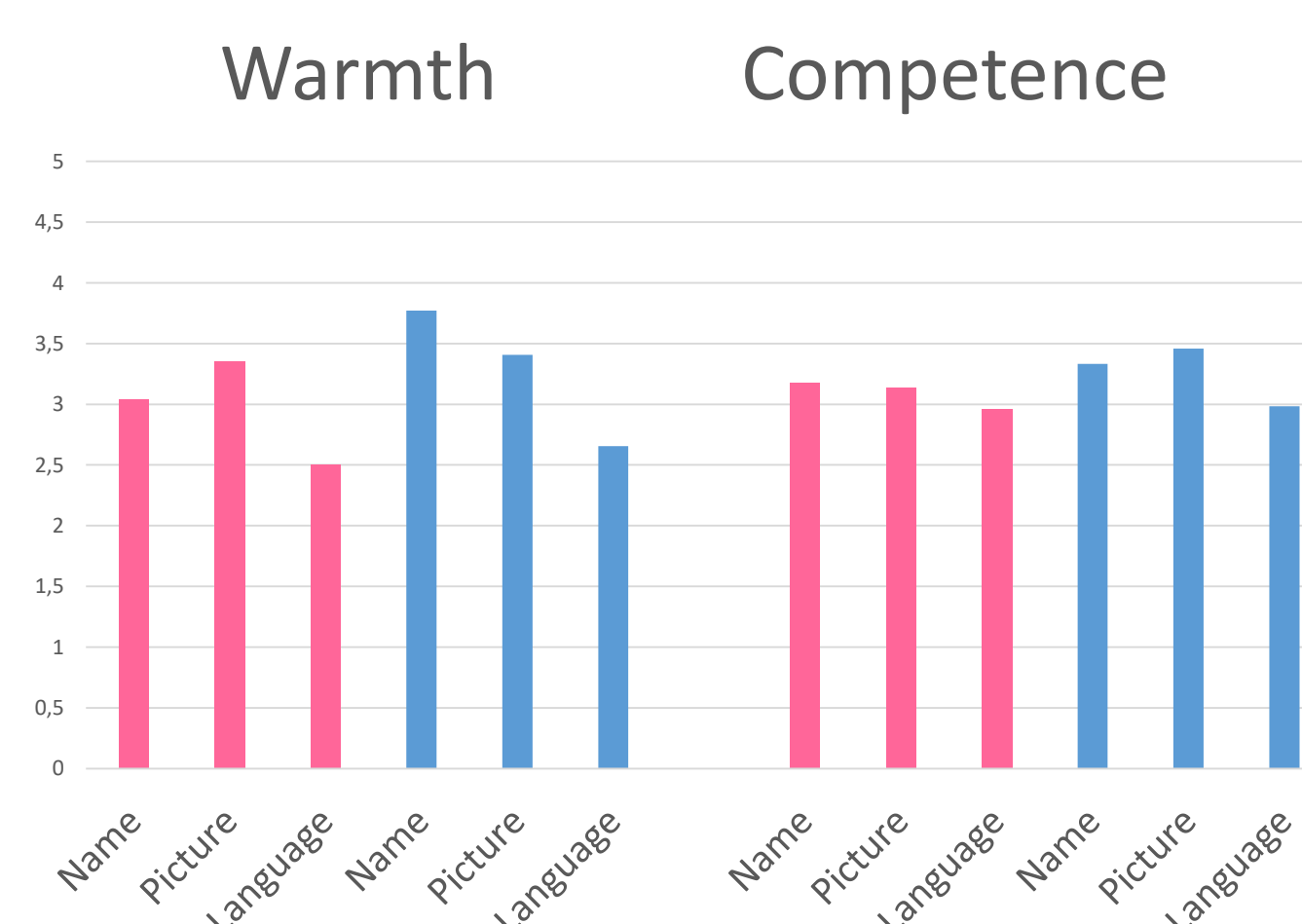


Illustration 3: Comparison of gender in dependency of various shaping

Added value

The studies' findings emphasize that gender stereotypes and their consequences are deeply rooted in the human's nature. Moreover, specific representation of the assistant's gender seem to boost the application of gender stereotypes.

References

- Eckes, T. (2002). Paternalistic and Envious Gender Stereotypes: Testing Predictions from the Stereotype Content Model. *Sex Roles : a Journal of Research*, 47(3/4), 99–114
- Lee, E. J., Nass, C., & Brave, S. (2000, April). Can computer-generated speech have gender?: an experimental test of gender stereotype. In *CHI'00 extended abstracts on Human factors in computing systems* (pp. 289-290). ACM.
- Nass, C., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates?. *International Journal of Human-Computer Studies*, 45(6), 669-678.
- Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, 308(5728), 1623-1626.