

Can Life-Like Virtual Interviewers Increase the Quality of Answers to Open Questions?

Höhne^{1,2}, Neuert³, & Claassen^{1,2}

¹ German Center for Higher Education Research and Science Studies (DZHW)

² Leibniz University Hannover

³ GESIS – Leibniz Institute for the Social Sciences

77th Annual WAPOR Conference

Seoul (South Korea) – July 28 to 31, 2024

Introduction I

- Demand for high-quality data from web surveys (Knowledge Sourcing Intelligence, 2023)
- Cost-efficient and streamlined web surveys replace other survey modes, especially in-person interviews (Schober, 2018)
 - *Even large-scale social surveys, such as ESS, start utilizing web surveys*
- Web surveys may not be good candidates for primary survey mode
 - *Depressed response rates* (Daikeler et al., 2020)
 - *Struggle with achieving high data quality* (Callegaro et al., 2015)
- Absence of interviewers impedes the provision of assistance and the creation of trust, motivation, and engagement

Introduction II

- Fusing elements of interviewer-based and web surveys
 - *Life-like virtual interviewers and self-administration*
- **Few studies utilized virtual interviewers** (Conrad et al., 2015; Conrad et al., 2020; Lind et al., 2013; Schuetzler et al., 2018)
 - *Limited knowledge about respondent satisfaction and data quality benefits*
 - *Frequently conducted in lab settings using closed instead of open questions*
- Open questions may place burden on respondents
 - *Especially, when answering through smartphones with on-screen keypad*
 - *Increased levels of item-nonresponse and short/incomplete answers*
- Following social interface theory
 - *Humanizing cues in a computer interface can change respondent behavior*
 - *Answers become similar to answers in human-to-human interaction* (Reeves & Nass 1997; Tourangeau et al., 2003)

Research Questions (RQs)

- Can life-like virtual interviewers increase answer quality of open questions in web surveys? (RQ1)
- How do life-like virtual interviewers affect respondents' web survey experience? (RQ2)
- How are life-like virtual interviewers evaluated by respondents? (RQ3)

Method: AI of Virtual Interviewers

Image generation

Instruction via descriptive keywords using diffusion models or architected transformers (Zhang et al., 2023)

Text generation

Generative Pretrained Transformers and Large Language Models (Vaswani et al., 2017)

Text-to-Speech generation

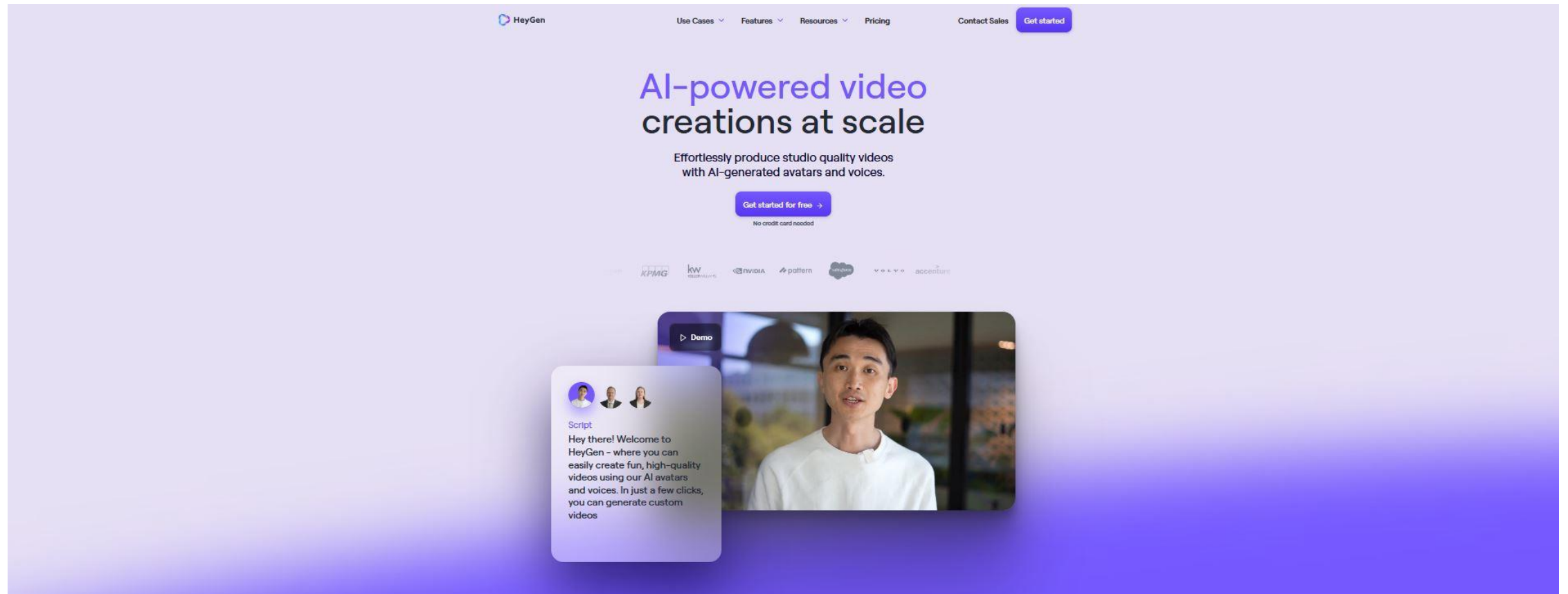
Statistical Parametric Synthesis and Neural Speech Synthesis (Tan et al., 2021): **1** “text analysis” (heteronyms), **2** “voice parameter prediction” (acoustic model), and **3** “vocoder analysis” (audio snippets)

Speech and image animation

Multistep pipelines of transformers, Recurrent Neural Networks, Convolutional Neural Networks, and Generative Adversarial Networks (Chen et al., 2023)

See Cheung, B. (2023): <https://bennycheung.github.io/create-personal-animated-ai-avatar>

Method: Creating Virtual Interviewers



See <https://www.heygen.com/>

Method: Study Design



Male casual
n = (376)




Male business casual
(n = 375)



Female casual
n = (395)



Female business casual
(n = 343)

- Experiment in a smartphone survey (N = 1,871)
- Between-subject design with 5 groups
 - *1 text control without virtual interviewer (n = 382)*
 - *4 treatment groups with a virtual interviewer*
- 2 open questions on women at the workplace (1) and family relations (1)
- Respondents had to click  for playing the virtual interviewer videos
 - *Respondents were informed that they are surveyed by virtual interviewers*
 - *Videos could be played multiple times*

Method: Example Question Screenshots

Könnten Sie bitte Ihre Meinung zu der Aussage „Wenn Frauen Probleme bei der Arbeit haben, übertreiben sie häufig.“ näher begründen?

Bitte geben Sie Ihre Antwort in das Textfeld ein.

Text control



Bitte geben Sie Ihre Antwort in das Textfeld ein.

Male casual



Bitte geben Sie Ihre Antwort in das Textfeld ein.

Female casual

Method: Sample

Data collection was conducted in the ResponDi/Bilendi panel in Germany in November and December 2023

Cross quotas: Age and gender plus quotas on education

Mean age: 49 years

Gender: 49% females

Education: 44% lower secondary school

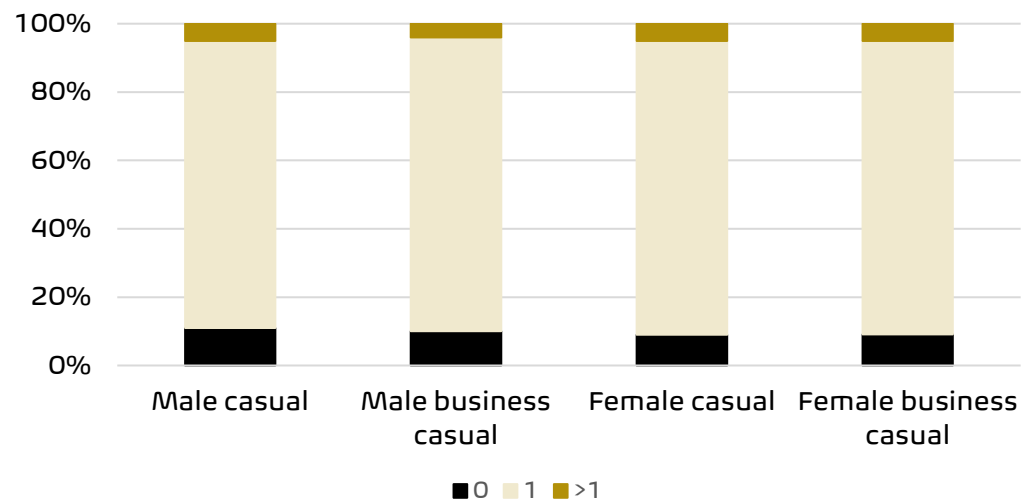
24% intermediate secondary school

34% at least college preparatory secondary school

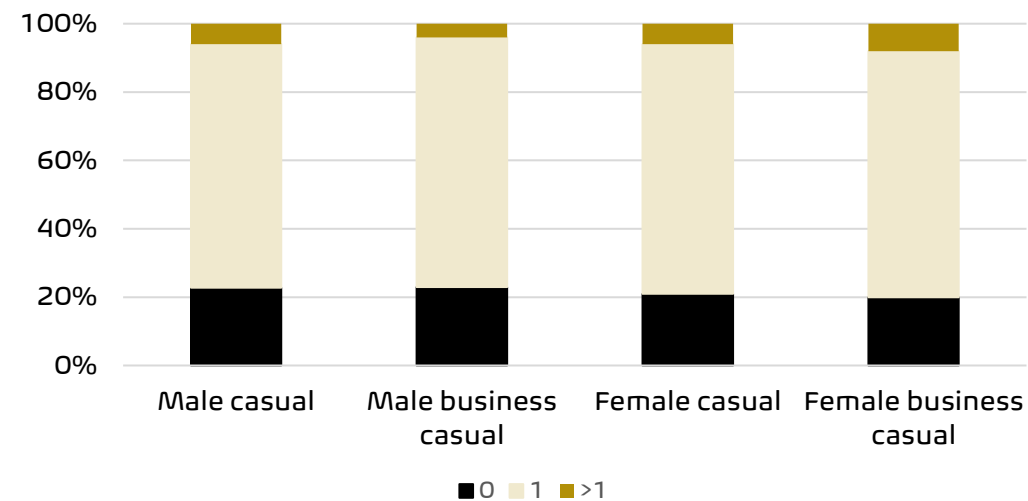
Experimental groups do not statistically differ with respect to age, gender, and education

Results: Video Play Rate

Open Question 1
Percentages

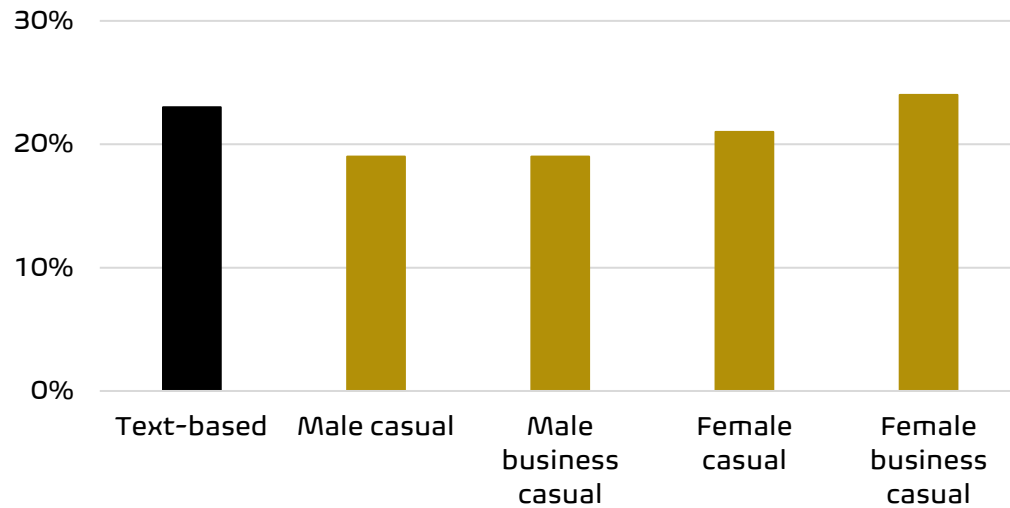


Open Question 2
Percentages



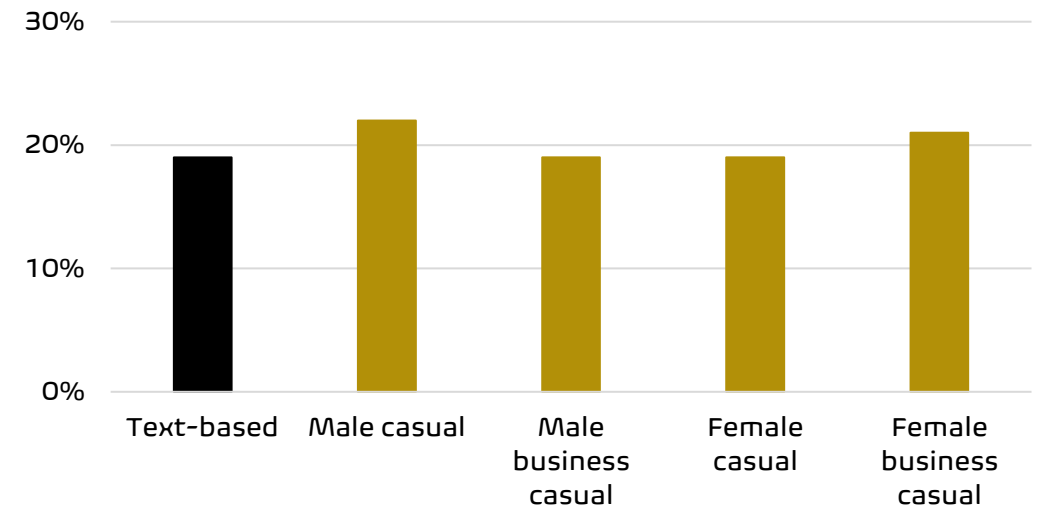
Results: Item-Nonresponse I

Open Question 1
Percentages



Chi-square test: $p > 0.05$. Item-nonresponse includes non-sense and non-substantive answers

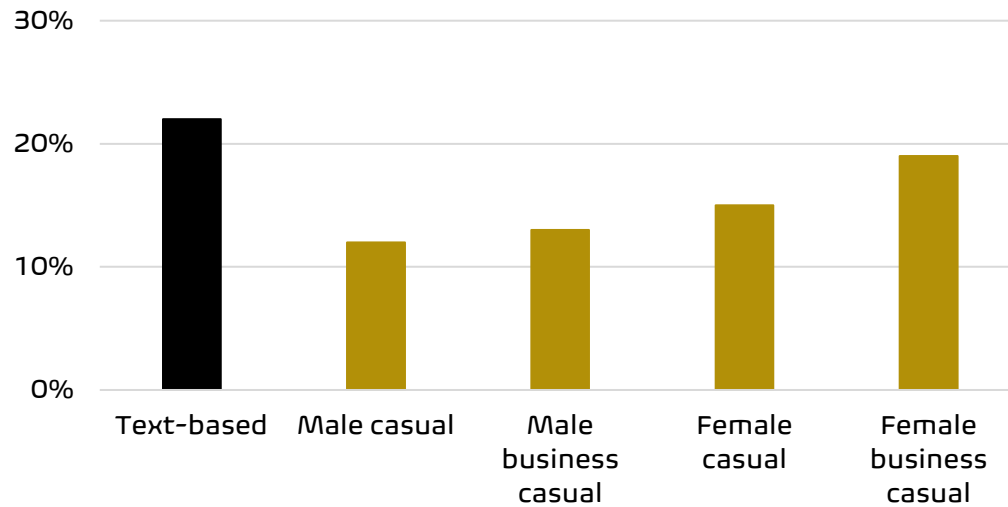
Open Question 2
Percentages



Chi-square test: $p > 0.05$. Item-nonresponse includes non-sense and non-substantive answers

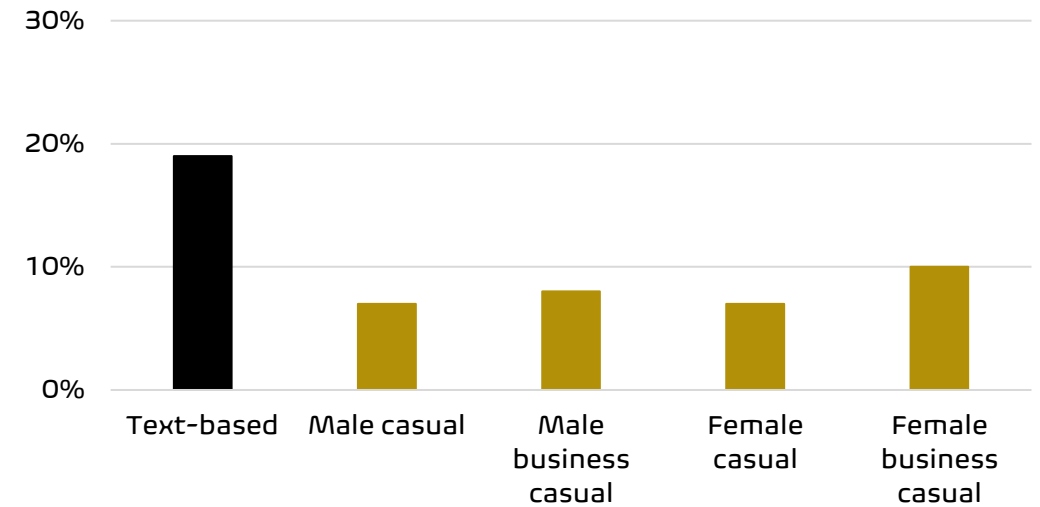
Results: Item-Nonresponse II

Open Question 1
Percentages



Chi-square test: $p < 0.05$. Item-nonresponse includes non-sense and non-substantive answers

Open Question 2
Percentages

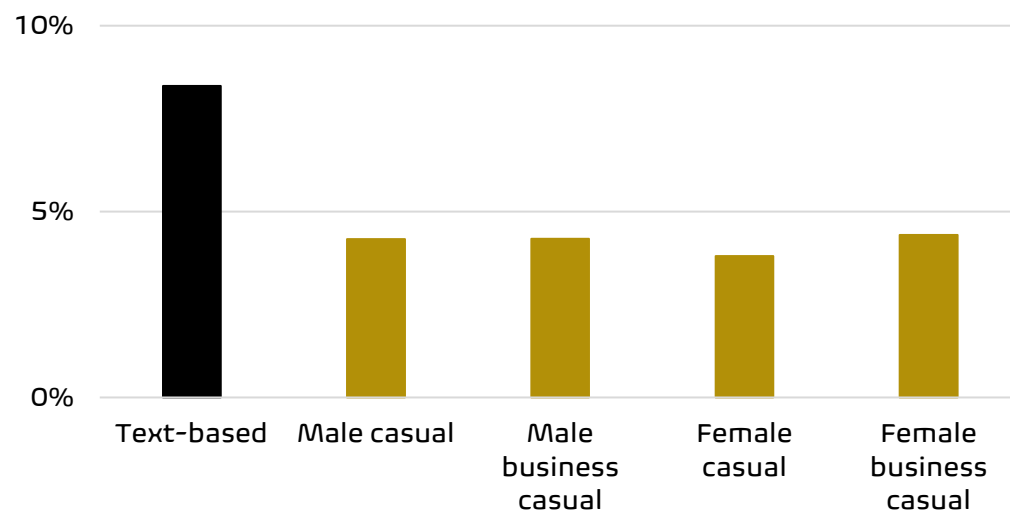


Chi-square test: $p < 0.05$. Item-nonresponse includes non-sense and non-substantive answers

Considering only "video playing" respondents
item-nonresponse is significantly lower

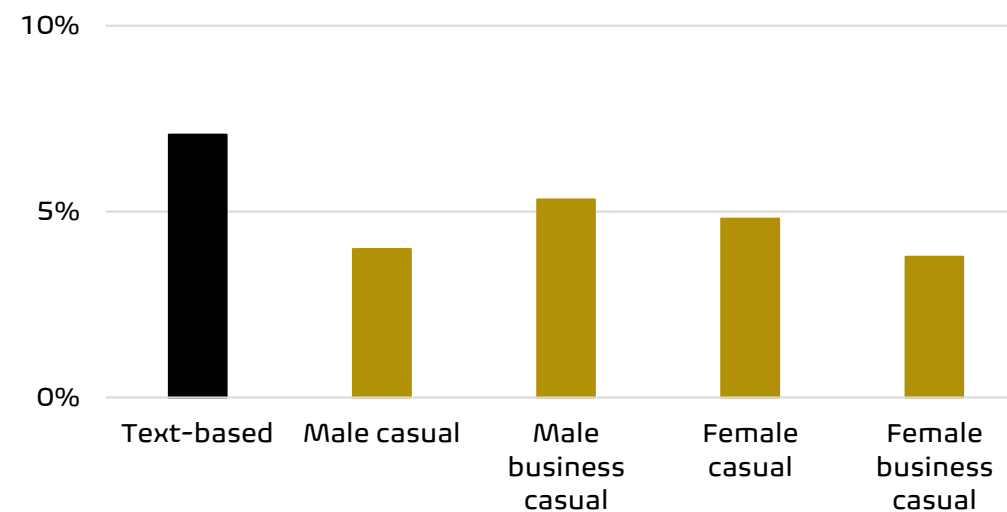
Results: Speeding

Open Question 1
Percentages



Chi-square test: $p < 0.05$. Fastest 5%-percentile

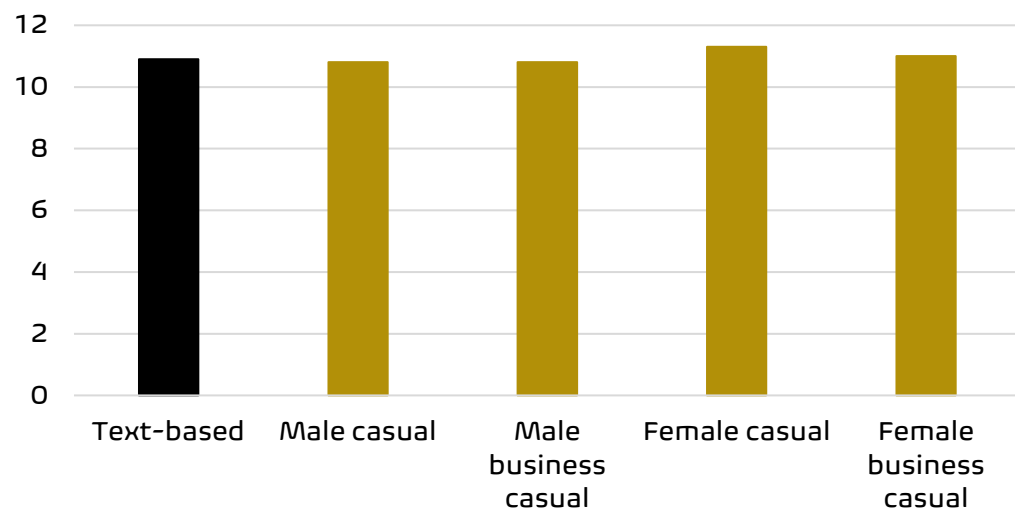
Open Question 2
Percentages



Chi-square test: $p > 0.05$. Fastest 5%-percentile

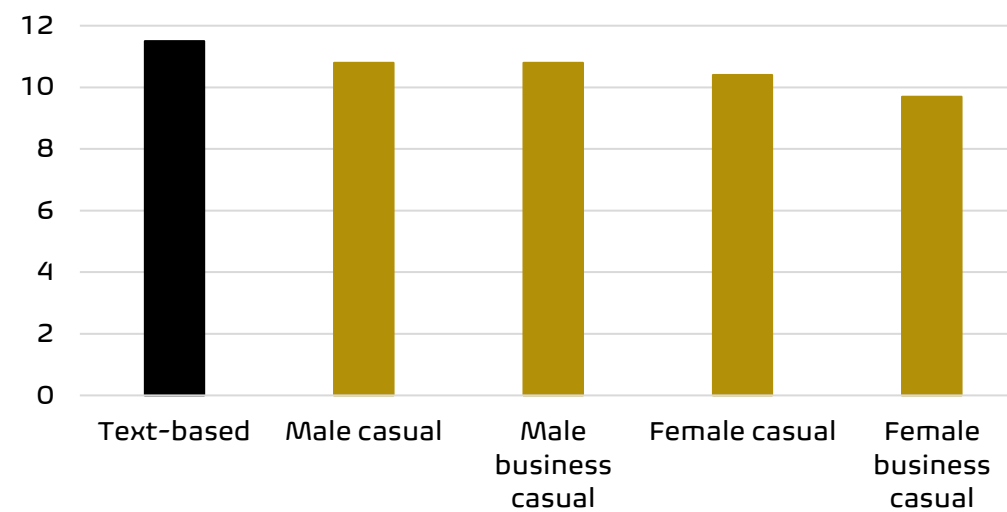
Results: Number of Words

Open Question 1
Average Count



One-way ANOVAs: $p > 0.05$

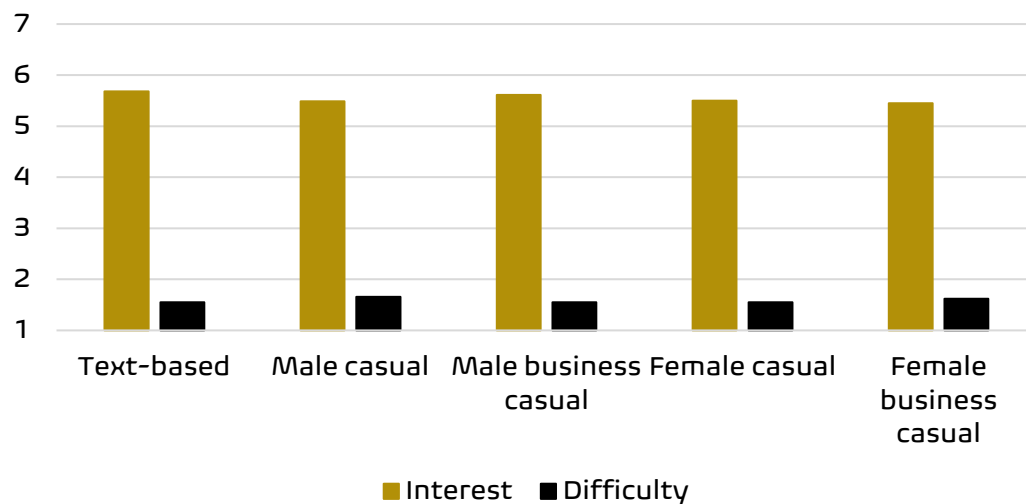
Open Question 2
Average Count



One-way ANOVAs: $p > 0.05$

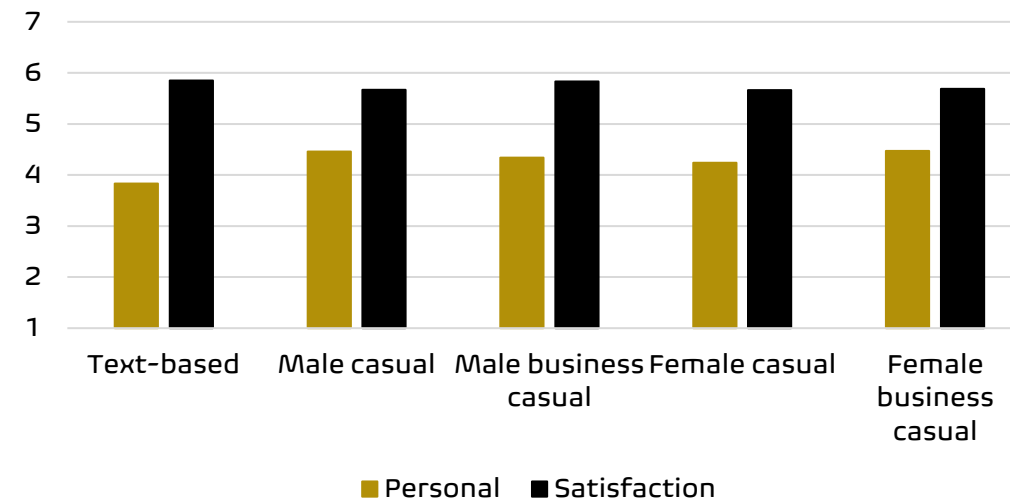
Results: Web Survey Experience

Survey evaluations
1 'Not at all' to 7 'Very'



One-way ANOVAs: $p > 0.05$

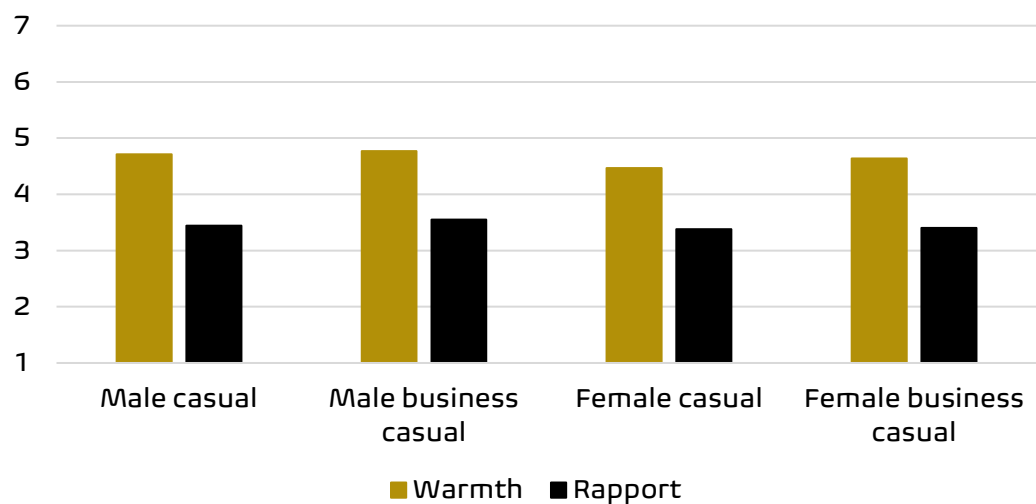
Survey evaluations
1 'Not at all' to 7 'Very'



One-way ANOVAs: $p < 0.05$ (only for personal)

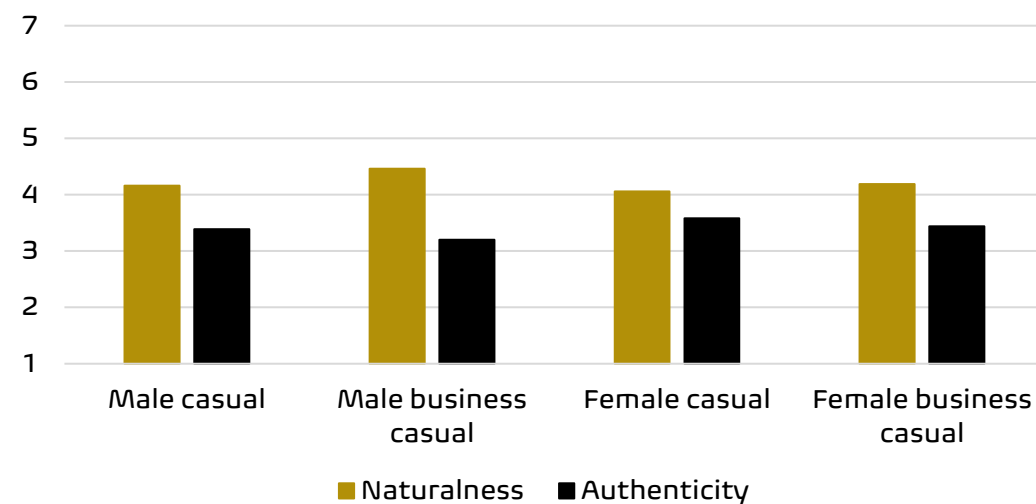
Results: Virtual Interviewer Evaluations

Interviewer evaluations
1 'Not at all' to 7 'Very'



One-way ANOVAs: $p > 0.05$

Interviewer evaluations
1 'Not at all' to 7 'Very'



One-way ANOVAs: $p > 0.05$

Discussion and Conclusion

- Item-nonresponse corresponds to prior studies on open questions
 - *Considering video playing reduces item-nonresponse in virtual interviewer conditions*
 - *About 85% of respondents comply with video playing*
- Speeding is more common in the text-based condition
 - *Virtual interviewers slow down respondents*
- No difference in word count between conditions
 - *Smartphones with a virtual on-screen keypad shrink viewing space of survey content*
- All virtual interviewer conditions are evaluated as being more personal
 - *Interest and satisfaction ratings are high (> 5 on 7-point scales)*
- Virtual interviewers perform similarly well in terms of respondent evaluations
- Take home message: Virtual interviewers have answer quality benefits and are evaluated well by respondents

Many thanks for your attention!

www.jkhoehne.eu
@jkhoehne

Literature I

- Callegaro, M., Lozar Manfreda, K., & Vehovar, V. (2015). *Web survey methodology*. Sage.
- Conrad, F. G., Schober, M. F., Hupp, A. L., West, B. T., Larsen, K. M., Ong, A. R., & Wang, T. (2023). Video in survey interviews: Effects on data quality and respondent experience. *Methods Data Anal.*, 17(2), 135-170.
- Conrad, F. G., Schober, M. F., Nielsen, D., & Reichert, H. (2020). Virtual interviewers, social identities, and survey measurement error. In K. Olson, J. D. Smyth, J. Dykema, A. L. Holbrook, F. Kreuter, & B. T. West (Eds.), *Interviewer effects from a total survey error perspective* (pp. 149–164). CRC Press.
- Conrad, F. G., Schober, M. F., Jans, M., Orłowski, R. A., Nielsen, D., & Levenstein, R. (2015). Comprehension and engagement in survey interviews with virtual agents. *Front. Psychol.*, 6, 1578.
- Daikeler, J., Bosnjak, M., & Lozar Manfreda, K. (2020). Web versus other survey modes: An updated and extended meta-analysis comparing response rates. *J. Surv. Stat. Methodol.*, 8(3), 513–539.
- Knowledge Sourcing Intelligence. (2023). Global online survey software market size, share, opportunities, COVID 19 impact, and trends by application, by product, and by geography – forecasts from 2023 to 2028. <https://www.knowledge-sourcing.com/report/global-online-survey-software-market>
- Lind, L. H., Schober, M. F., Conrad, F. G., & Reichert, H. (2013). Why do survey respondents disclose more when computers ask the questions? *Public Opin. Q.*, 77(4), 888–935.
- Reeves, B., & Nass, C. (1997). *The media equation: how people treat computers, television, and new media like real people and places*. Cambridge: CSLI and Cambridge University Press.
- Schober, M. F. (2018). The future of face-to-face interviewing. *Qual. Assur. Educ.*, 26(2), 290–302.
- Schuetzler, R. M., Giboney, J. S., Grimes, G. M., & Nunamaker, J. F. (2018). The influence of conversational agent embodiment and conversational relevance on socially desirable responding. *Decis. Support. Syst.*, 114, 94-102.

Literature II

- Tourangeau, R., Couper, M. P., & Steiger, D. M. (2003). Humanizing self-administered surveys: Experiments on social presence in web and IVR surveys. *Computers in Human Behavior*, 19, 1–24.
- West, B. T., Ong, A. R., Conrad, F. G., Schober, M. F., Larsen, K. M., & Hupp, A. L. (2022). Interviewer effects in live video and prerecorded video interviewing. *J. Surv. Stat. Methodol.*, 10(2), 317–336.
- Zhang, C., & Conrad, F. G. (2018). Intervening to reduce satisficing behaviors in web surveys: Evidence from two experiments on how it works. *Social Science Computer Review*, 36(1), 57–81.

Appendix: Question Translations

- Open Question 1:
 - *Please explain your opinion on the statement 'When women have problems at work, they often exaggerate' in more detail.*
- Open Question 2:
 - *In your opinion, what is the ideal division of labor between men and women concerning work and family?*