

How to incorporate AI interviewers in contemporary web surveys?

Höhne^{1,2}, Broneske², Neuert³, & Claassen^{1,2}

¹ Leibniz University Hannover

² German Center for Higher Education Research and Science Studies (DZHW)

³ GESIS – Leibniz Institute for the Social Sciences

Online Event of the Royal Statistical Society

10th January 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 are not ready for taking over
 - *Depressed response rates* (Daikeler et al., 2020)
 - *Impede participation for illiterate people* (Höhne, 2023)
 - *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

- Advances in communication technology and AI introduce new data collection opportunities
- Fusing elements of interviewer-based and web surveys
 - *Life-like AI 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)
 - *Inconclusive results on respondent satisfaction and data quality*
 - *Frequently conducted in lab settings*
- **Web surveys with pre-recorded interviewers have quality benefits** (West et al., 2022; Conrad et al., 2023)
 - *More disclosure of sensitive behavior, less rounding, and less error variance*

Research Questions (RQs)

- How ...

... are AI interviewers evaluated by respondents? (RQ1)

... do AI interviewers affect respondents' survey satisfaction? (RQ2)

... do AI interviewers affect response effort in terms of response times? (RQ3)

... do AI interviewers affect question reading (or video playing)? (RQ4)

... do AI interviewers affect item-nonresponse? (RQ5)

What is the AI Part of AI 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

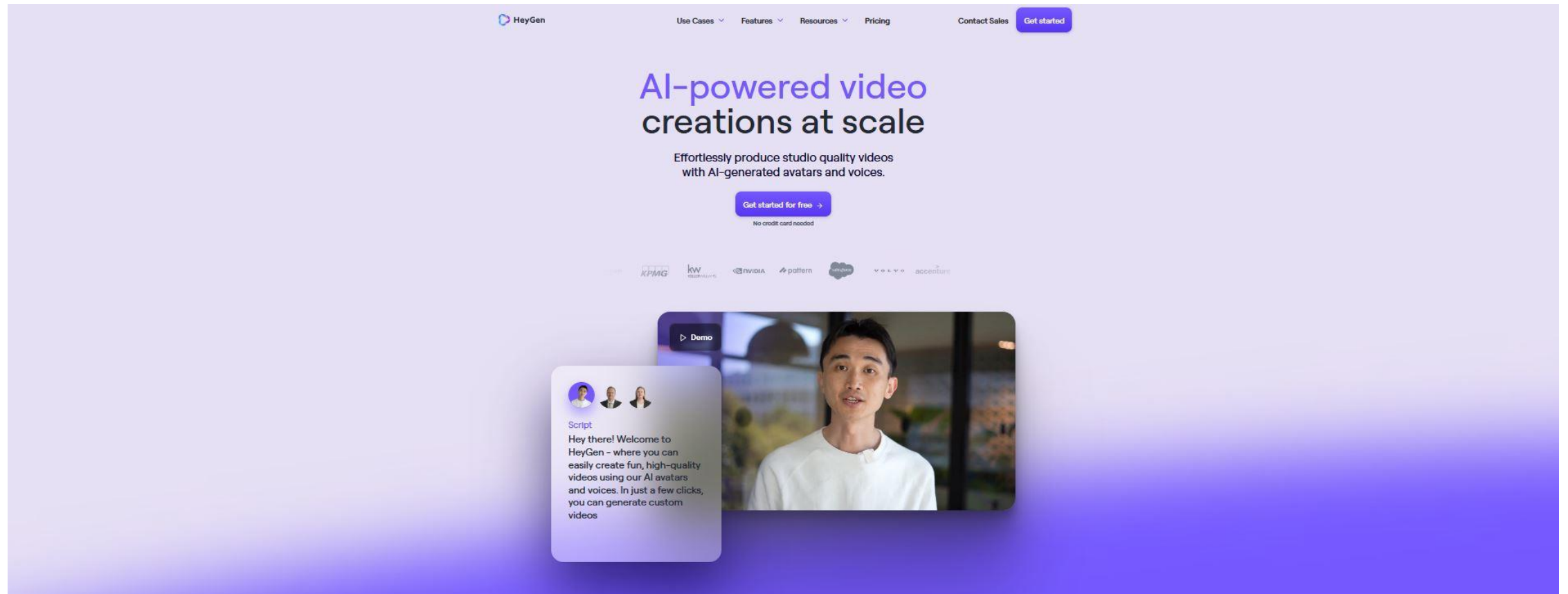
Statistical Parametric Synthesis and Neural Speech Synthesis (Tan et al., 2021): **1** “text analysis” (heteronyms), **2** “voice parameter prediction” (acoustic model), and **3** “vocoder analysis” (audio snippets)

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 AI 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 = 1,489)
 - *Response time collection via ECSP* (Schlosser & Höhne, 2020)
- Between-subject (4-group) design
 - *AI interviewers were called Alex*
- Closed and open questions
 - *6 closed and 2 open questions (e.g., on family relations)*
- Respondents had to click  for playing the AI interviewer video
- Questions for evaluating AI interviewers and survey satisfaction
 - *Placed at the end of the survey*

Method: Sample

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

Cross quotas: Age and gender plus education

Mean age: 49 years

Gender: 49% females

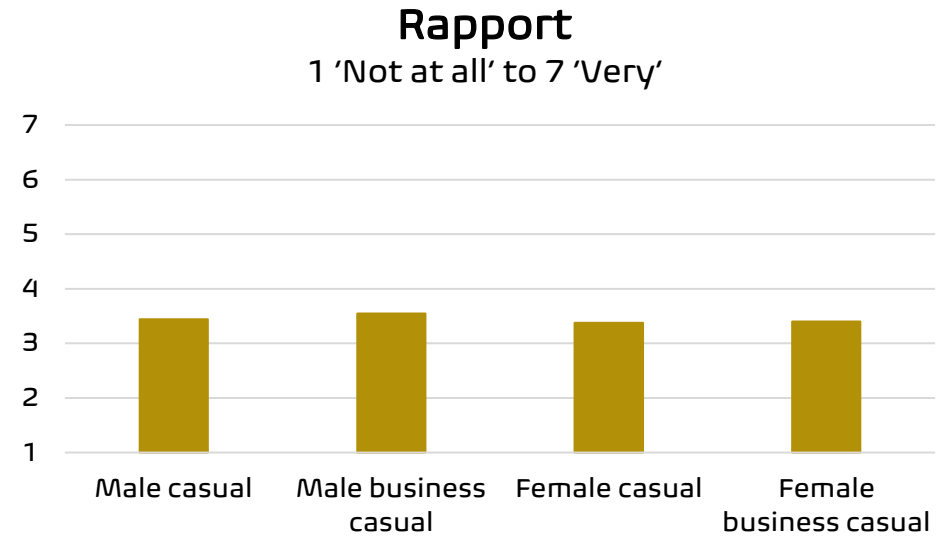
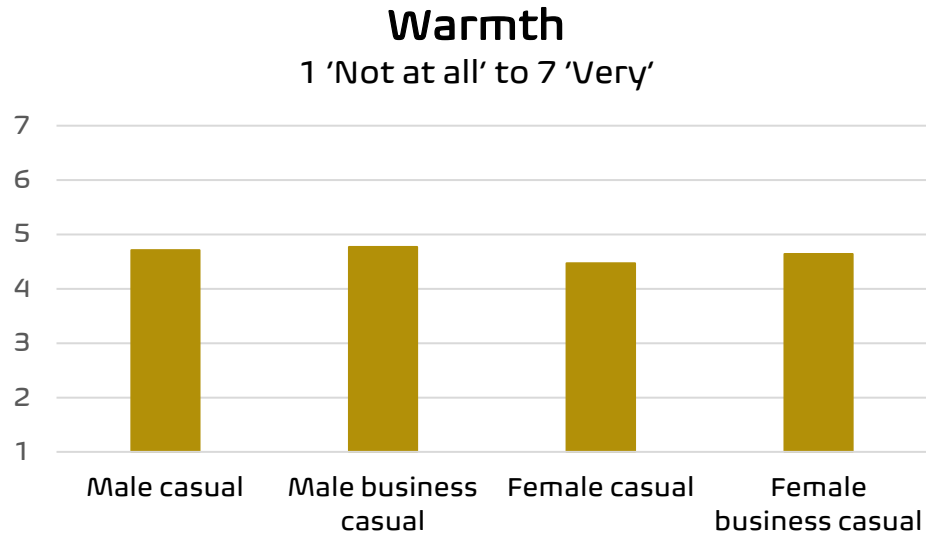
Education: 43% lower secondary school

24% intermediate secondary school

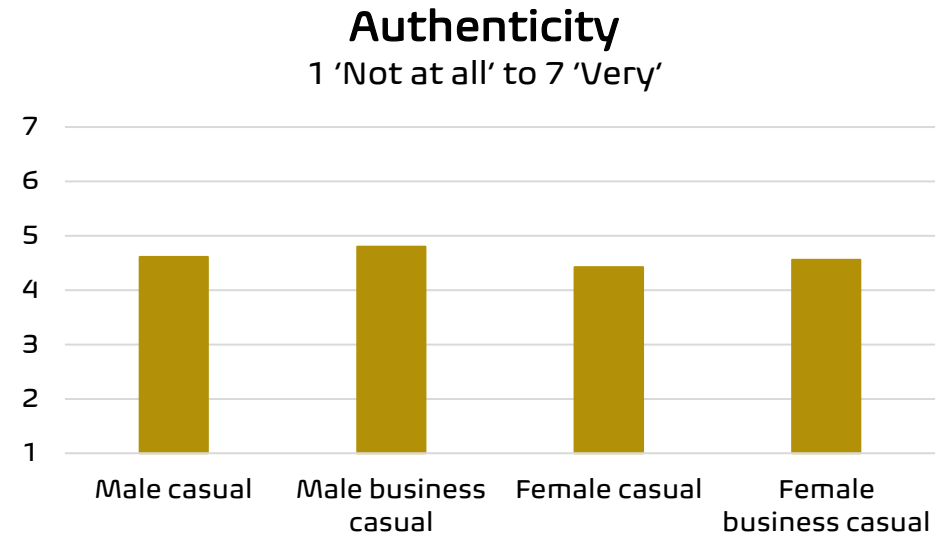
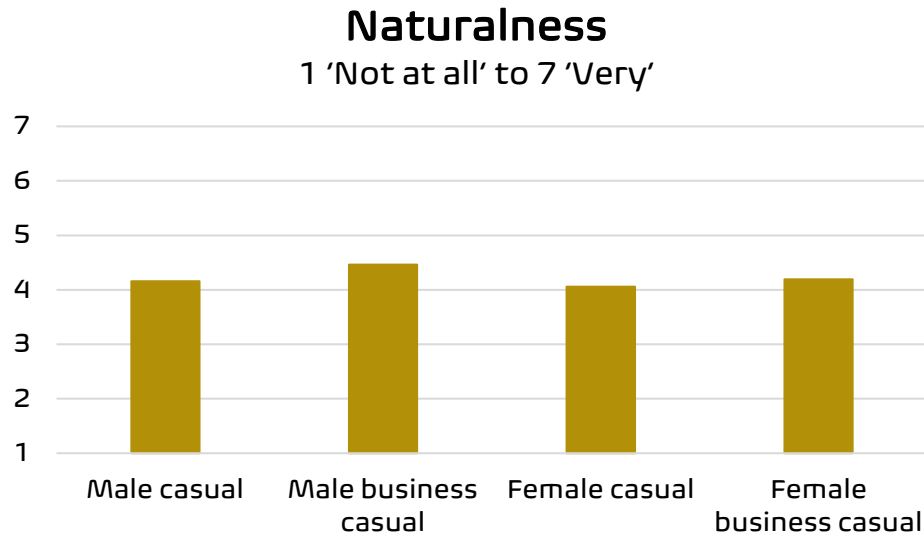
33% at least college preparatory secondary school

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

Results: Research Question 1

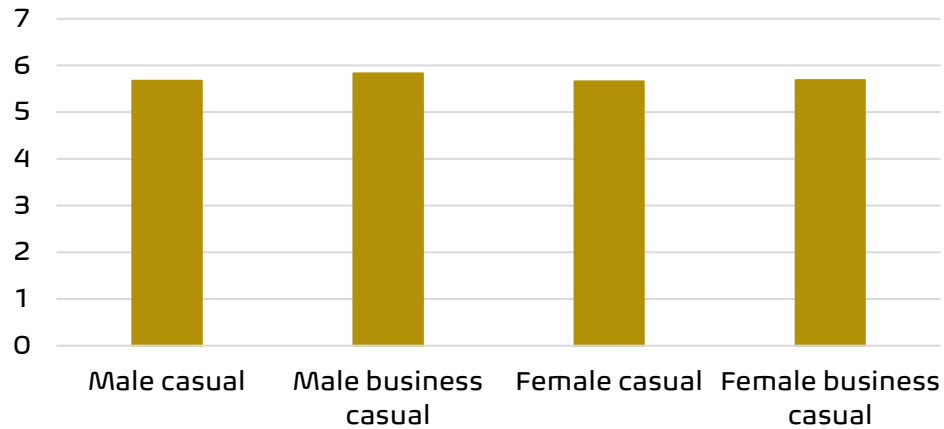


Results: Research Question 1

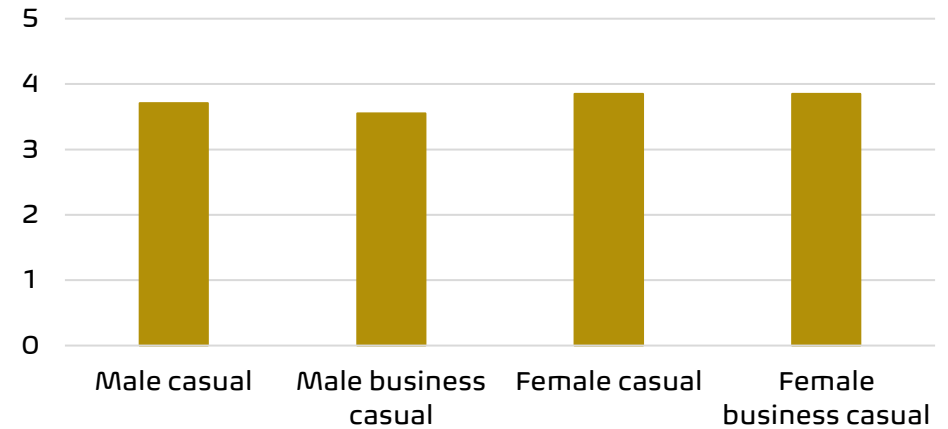


Results: Research Questions 2 & 3

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

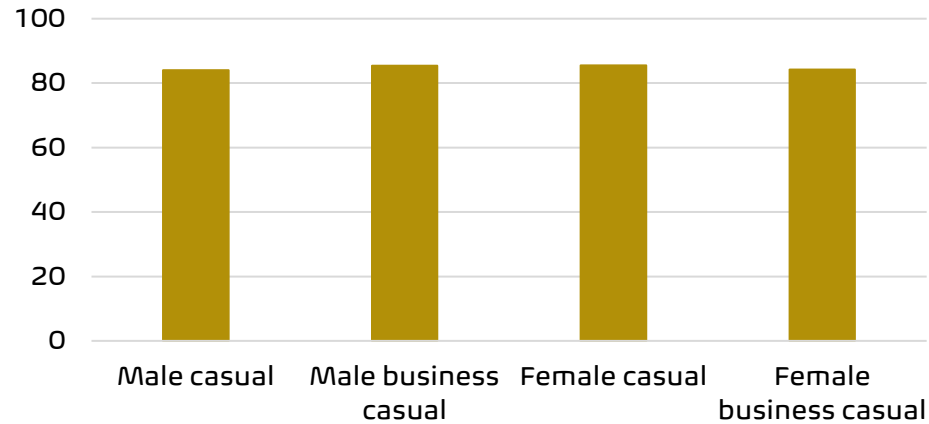


Response time
Minutes

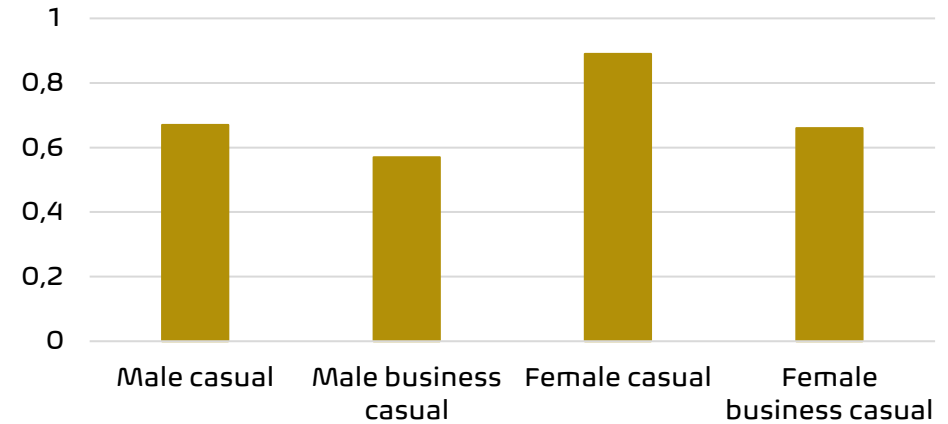


Results: Research Questions 4 & 5

Video Playing (at least once)
Percentages



Item-nonresponse
Percentages



Discussion and Conclusion

- AI interviewers perform similarly well in terms of respondent evaluations
 - *Rapport evaluations are somewhat lower than the remaining evaluations*
- There is no difference regarding survey satisfaction
 - *However, survey satisfaction appears to be high*
- AI interviewers pose similar response effort
 - *Different speech behaviors do not introduce comprehension or answer difficulties*
- Importantly, majority of respondents plays the videos (> 80%)
 - *Compliance with survey task and general survey question processing*
- No increase in item-nonresponse across AI interviewers
 - *Item-nonresponse is slightly higher (~2%) for open questions*

Future Avenues

- Conducting subgroup analysis
- Comparing data quality to text-based web surveys
- Including more natural answer formats
 - *For example, voice communication channels*
- Growing beyond commercial providers
 - *Developing open-source infrastructure for AI interviewer web surveys*
 - *Tailoring AI interviewers to survey purposes*
- Moving from avatars to agents
 - *Responsive and autonomous AI interviewers*

Many thanks for your attention!

www.jkhoehne.eu
@jkhoehne

Literature I

- Callegaro, M., Lozar Manfreda, K., & Vehovar, V. (2015). *Web survey methodology*. Sage.
- Chen, L., Wu, Z., Li, R., Bao, W., Ling, J., Tan, X., & Zhao, S. (2023). VAST: Vivify your talking avatar via zero-shot expressive facial style transfer. *arXiv*. DOI: 10.48550/arXiv.2308.04830
- 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. DOI: 10.12758/mda.2022.13
- 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. DOI: 10.3389/fpsyg.2015.01578
- 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. DOI: 10.1093/jssam/smz008
- Höhne, J. K. (2023). Are respondents ready for audio and voice communication channels in online surveys? *Int. J. Soc. Res. Methodol.*, 26(3), 335-342. DOI: 10.1080/13645579.2021.1987121
- 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. DOI: 10.1093/poq/nft038
- Schlosser, S., & Höhne, J. K. (2020). ECSP – Embedded Client Side Paradata. Zenodo. DOI: 10.5281/zenodo.3782592

Literature II

- Schober, M. F. (2018). The future of face-to-face interviewing. *Qual. Assur. Educ.*, 26(2), 290–302. DOI: 10.1108/QAE-06-2017-0033
- 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. DOI: 10.1016/j.dss.2018.08.011
- Tan, X., Qin, T., Soong, F., & Liu, T. Y. (2021). A survey on neural speech synthesis. *arXiv*. DOI: 10.48550/arXiv.2106.15561
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In *NIPS'17*. Curran Associates, Inc. <https://dl.acm.org/doi/10.5555/3295222.3295349>
- 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. DOI: 10.1093/jssam/smab040
- Zhang, C., Zhang, C., Zhang, M., & Kweon, S. (2023). Text-to-image diffusion models in generative AI: A survey. *arXiv*. DOI: 10.48550/arXiv.2303.07909