How do Conversational Agents Transform Qualitative Interviews? Exploration and Support of Researchers' Needs in Interviews at Scale

Bulygin, Denis

DOI
10.1145/3490100.3516478

Publication date
2022

Document Version
Final published version

Published in
27th International Conference on Intelligent User Interfaces, IUI 2022 Companion

How do Conversational Agents Transform Qualitative Interviews? Exploration and Support of Researchers' Needs in Interviews at Scale

Denis Bulygin
d.bulygin@tudelft.nl
Delft University of Technology
Delft, Netherlands

ABSTRACT

In recent years, conversational agents (CAs) have been receiving more attention as tools for collecting data through qualitative interviews. The problem is we know little about how CAs affect both the interviewees and interviewers. This PhD project is dedicated to studying how to evaluate CA-mediated interviews and their effects on participants (both interviewees and interviewers). The findings of this project will allow us to support the interview practitioners with the tools for interview analytics and interview data analysis. It will be especially helpful in the large-scale settings which CA-mediated interviews enable. This proposal describes State-of-the-art on the topic and presents the motivation of a study with key research questions to answer.

CCS CONCEPTS

• **Information systems → Information systems applications**; • **Human-centered computing → Interaction design process and methods**.

KEYWORDS

interviews, large-scale interviewing, conversational agents, chatbots

ACM Reference Format:
Denis Bulygin. 2022. How do Conversational Agents Transform Qualitative Interviews? Exploration and Support of Researchers' Needs in Interviews at Scale. In *27th International Conference on Intelligent User Interfaces (IUI '22 Companion), March 22–25, 2022, Helsinki, Finland.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3490100.3516478

1 INTRODUCTION

This paper presents a plan on studying the effects of interview Conversational Agents (CAs) on its participants. Interviewing is a qualitative method that is widely used (although, not limited to) by designers to understand the users with their thoughts, experiences, and emotions. It helps when one explores the contexts of supported activity, generates requirements, and evaluates the product [13].
Recently, Conversational AI (CAI) has emerged in many areas of society and business in the form of CA. Prior work presents examples of CAs that conduct structured [30], cognitive [16], and job [32] interviews. Currently, most of the interviews are conducted by a human interviewer with a small number of participants which takes time and does not easily scale up. The adoption of CA-mediated interviews will allow us to gather data automatically which will help us:

(1) To gather larger bodies of data. It will bring quantitative approaches to interview settings. Currently, large-scale interview studies are resource-intensive and time-consuming, both during the data collection and data analysis stages. This transformation will enable us to come up with more reliable research findings.
(2) To conduct experiments with the methodological questions of qualitative interviewing. There are a lot of guidelines mostly based on practical experience and theoretical reasoning. CA-mediated interviewing will facilitate testing them and understanding their effects statistically.
(3) To reach a wider audience. If interviews can be outsourced to a CA, the researchers will not need to think about time zones and will be able to recruit more people around the globe.

However, we know little about how CAs change interviewees during a conversation and how it affects the data quality. We have yet to find out how researchers will change the practice of interviewing when CA-mediated interviewing is mass adopted. Otherwise, we will use systems in situations we poorly understand and will answer questions or act upon data with low credibility. That's why we must explore the effects of CA's on the human actors from both sides of interviewing: interviewees and interviewers.
This PhD project aims to contribute to CA-mediated interviewing in two ways: 1) To study how to evaluate interviews and highlight methodological problems in CA-mediated interviews; 2) To analyze how CA-mediated interviews transform research practices and to design the tools that support researchers in such transformation. As a result, the main beneficiaries of this work will be the designers of interview CAI and roles that actively use interviewing in their practice.

2 RELATED WORKS

2.1 Overview of the CAs in qualitative interviewing

The first conversational agent (CA) was developed half a century ago and was called Eliza [28]. It was a social CA that could talk with people on a variety of topics. 15 years ago Elizabeth was presented [11] – a task-oriented CA whose purpose was to assist humans in

IUI '22 Companion, March 22–25, 2022, Helsinki, Finland

ACM ISBN 978-1-4503-9145-0/22/03.
https://doi.org/10.1145/3490100.3516478

qualitative research. Ten years ago [18] presented the first embodied agent which combines conversational interface with a bunch of other input tools (e.g. camera) to gather more information about the interviewee. These initial studies, although with their limitations, showed the promise in using conversational agents for data gathering and shaped the way we design CA (e.g. how to evaluate CA during an interview? [11]).

Recent studies explore the gamut of applications of CAs in qualitative interviewing: fully structured research interviews [8, 30], job interviews [32], cognitive interviews [16], diary studies [2, 4, 15, 17]. The group of researchers [8, 29, 30] have built a platform Juji.io. This platform enables the interviewers to make fully structured interviews with the capability to ask follow-up questions [30] and paraphrase user input [29]. The same platform can be used in job interviews [32] and allows experimenting with CA's personality. Platform also helps to learn how to support interview designers by analyzing conversation and suggesting improvements of the interview guide [8].

Minhas et al. [16] has built a CA to automatically conduct cognitive interviews. CAs can be programmed to ask questions in a particular way, which is vital in cognitive interviewing; however, it was unclear how participants would react to a CA. The authors conducted a study and interviewees showed no negative reaction towards CA. Some authors also apply CAs in diary-based study settings [2, 15, 17] and surveys [10] which ensures methodological triangulation with interviews. CAs help to collect the data using text-based [2] or a voice-based [17] input or the mix of both [17]. The last authors built a prototype to assess how a voice-interface system would prove itself to capture hands-free experiences. The system gathers physical movements and records subjective reflections of humans using voice input. Mairittha et al. [15] used CA to record participant activities through voice utterances and compared text and voice inputs. Chang et al. [4] used CA to help the close ones of participants to supplement and comment on the gathered data during experience sampling. To sum up, CAs are used in a variety of research settings and show promise to become a new research tool for user studies.

2.2 Advantages of CAs in qualitative interviewing

CAs can work without communication fatigue. They allow the collection of data from a wider audience from any part of the world. Moreover, CAs can simultaneously communicate with many interviewees at the same time. Zhou et al. [32] have created CA to conduct job interviews and summarize key information in the dashboard visualizations. The hiring manager could assess the personality traits obtained from the conversation and reach responses to the questions for each participant. In this way, CA facilitated the hiring process and allowed the human to process 290 applicants in two hours. A human with no support from CA processed 400 applicants in two weeks.

CAs can follow precisely a predefined script and phrasing of the questions. Following the script is important in some settings. For example, cognitive interviews are used in criminology to reconstruct memories of accidents and criminal situations. It is vital to ask the questions as correctly as possible so that respondents do not transform their memories in the process. Sometimes following the phrasing is a challenge for humans. Minhas et al. [16] used a chatbot in a cognitive interview to avoid deviation from the script and question phrasing and it was a successful attempt.

People are more open to CAs. One of the great challenges in interviewing is making people trust an interviewer to share personal and sensitive information. Research shows people can be more open to CAs [10] which is helpful for interviews on sensitive topics. The findings are consistent with another study [27] that shows that people are more honest in impersonal communication: a conversation with a system or a conversation with the human through the text.

2.3 Challenges for CAs in qualitative interviewing

We have to teach CAs to conduct flexible formats of interviewing. Presented studies focus on fully-structured interviews which allows little to no digression from the main script (e.g. with paraphrasing or follow-up questions [29, 30]). Flexible formats, such as semi-structured interviews, are more difficult to navigate and manage. That's why training CAs for it is a challenging task that will take time for developers to figure out.

We have to teach CAs how to conduct interviews like an expert. Interviewers rely on their experience and best practices when making decisions about interviews. Unlike humans, CAs must be programmed beforehand to make such decisions. Before we can program interviewers' decisions, we need to formalize it in the set of rules which is yet to be done. For example, how does one understand if a respondent gives inconsistent or insincere responses?

CAs can affect the interview participants in unknown ways. Research shows the difference in data gathered by human and automatic systems [10, 26]: more disclosure, less humane thought expression from participants; however, these are the measured effects on survey data. Qualitative interviews produce different data which requires different approaches for analysis and different metrics of data quality.

CAs replacing researchers in the conversation can affect the researchers in unknown ways. Currently, interviewers interact with participants and control the flow and outcomes of the conversation. That's why they usually have a great understanding of the interview context which is especially helpful during analysis. Delegating parts of the interviews to CA can influence how researchers approach later steps of a study: they can observe more non-verbal details (if present in conversation) or they can analyze unfamiliar data (if absent in conversation). Either way, it will change the quality of the analysis and it is important to keep these effects in mind.

2.4 Research Gap and the scope of this PhD project

This PhD project is focused to overcome the last two last challenges: the lack of understanding about how CAs affect the interviewees and interviewers. Current research of CAs in interviewing is mostly focused on building better systems. However, we know little about how this new technology transforms the procedure. Research in the area of surveys shows the influence of both the automatic systems (vs humans) [26] and conversational interfaces (vs traditional forms) [10]. It is possible other conversational methods will alter as well. This project aims to explore the effects of conversational interaction on the procedure of the interview. The results of this project will be useful in several ways. It will help to make informed decisions about the automation of the interview process and facilitation of the interviewer's work. It will also help to build the tools supporting the work of the experts in the area of large-scale interviews. Lastly, it will help to connect the research in the methodology of interviews and conversational AI, allowing to conceptualize and conduct large-scale experiments to answer methodological questions.

3 RESEARCH QUESTIONS

As was stated earlier, the project is focused on the effects of CAs on interviewees and interviewers. That's why the project is aimed to explore two research themes (Fig. 1):

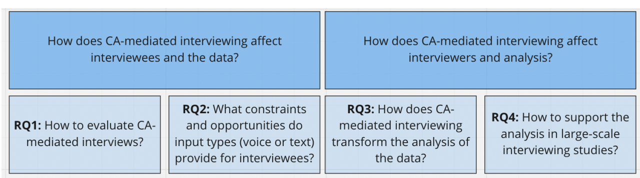

Figure 1: Research themes and Research Questions

3.1 How does CA-mediated interviewing affect interviewees and the data?

3.1.1 ***RQ1: How to evaluate CA-mediated interviews?*** Developers and interviewers will have to make decisions about the quality of CAs. When interviews are delegated to CAs, humans will not be involved in the process as much as they do in manual procedures. But they still need to detect problems and make adjustments in the study or a CA. If CAs provide affordances for quick and meaningful evaluation of interviews, it will allow stakeholders to ensure the methodological quality of the procedure and to improve systems much quicker.

With a few exceptions, prior research does not provide us with detailed guidelines on how to evaluate the data quality of interviews in general. Methodological research mostly focuses on the procedure [24]. For example, Rubin and Rubin [25] highlight "credibility" and "thoroughness". The closest one to the evaluation of the data is Kvale [12] who suggests assessing interviews based on the richness and length of interviewee responses among other things. Research of interview CA, authors mostly focus on proxy metrics such as non-response or IDK-response rate [10]. Alternatively, they use post-test questionnaires [2, 10] to capture respondents' reflections on the procedure.

Some authors [8] analyze the content of participants' responses to evaluate the effectiveness of interview CA. Authors have made a set of metrics extracted from conversations: response length, answer informativeness, engagement duration. Generally, it is accepted to aim for longer answers rich with meaningful details [12]. Measurement of these characteristics helps to evaluate the quality of particular questions [8] and can help to evaluate interviews as a whole.

Although these metrics are a step towards solving this problem, there are important measurements beyond these metrics that need to be taken. Interviewers struggle with social desirability, lack of disclosure, inconsistency of answers, memory biases. **My goal is to try to formalize the signs of these problems so that systems could provide stakeholders with meaningful information.**

3.1.2 ***RQ2: What constraints and opportunities do input types (voice or text) provide for interviewees?*** When RQ1 is answered, I will need to test generated metrics in the study. Currently, the goal is to apply generated metrics to explore the effects of voice- or text-based input on the interviewees. Interviewees interact with CAs using a technology which both supports textual and verbal input. While CAs can recognize human speech, it is easier to build textual input and avoid errors in understanding utterances [32]. Currently, we know little about this question but they may differentiate answers drastically.

Verbalizing can lead to longer and richer but to a less structured and less coherent answers because of the spontaneity of speech. Also, the inexperienced users may face problems [22] with voice interaction or be in an inappropriate context for answering with voice [10]. All these factors can result in lower response rates and data of lower quality [21].

Text typing is more user-friendly because more people are accustomed to it and it is a more private type of interaction. Also, in survey interviews, participants disclose sensitive information and give precise answers more often in text rather than in voice communication [27]. However, typing answers can be burdensome for interviewees. It takes more time to type the text than to answer vocally [9, 26]. It also requires more effort to think about the answer and phrase it properly in written form. Moreover, typing makes it difficult to talk with people within a particular context (e.g physical movements [17]), which should be less of a problem with verbal communication [7].

That is why we need to study how the input type affects interviewees and the quality of gathered data in the interviews. Prior work compares voice- and text- input in a data-gathering process [22, 26] but it does so in the settings of surveys and it focuses on closed-ended questions. I want to focus on studying these effects in interviews with open-ended questions and applied metrics developed in RQ1 to learn how the input type affects the gathered data and explore situations when one format is preferred over another.

3.2 How does CA-mediated interviewing affect interviewers and analysis?

3.2.1 ***RQ3: How does CA-mediated interviewing transform the analysis of the data?*** CA-mediated interviews will affect how human interviewers approach analysis in their study. Instead of leading a conversation human interviewers will be either observing an interview or be absent. Both of these outcomes should change the way interviewers take notes and analyze data and it will create new problems to solve. Interviewers supported by CA will be able to observe more details which should enrich the data and insights during analysis.

However, in some settings, CA can take over interviews. It can result in interviewers being absent from the conversation and being unfamiliar with gathered data when analyzing it. In addition, non-verbal cues (e.g., facial expressions) can be poorly captured with an absent human interviewer. Non-verbal cues play an important role in human communication, and thus in an interview too. During manual interviews, the interviewers may either pay attention to non-verbal cues and take notes or entirely lose this information, unless of course the interview is being recorded. When it comes to CAI interviewing, it is a challenge to capture and recognize such cues unless the human is taking notes or CA has appropriate hardware and software to do so [18].

3.2.2 ***RQ4: How to support the analysis in large-scale interviewing studies?*** CA-mediated interviews can change current analysis practices in two ways. First, it can help us solve the problems occurring in manual interviewing and analysis. Second, it will create new obstacles and challenges which we will need to overcome.

Regarding the problems with manual interviewing and analysis, two often discussed problems are personal bias of a researcher [13] and adhering to known theoretical scheme and categorization [5]. Not only do such problems lead to questioning research reliability but they also limit the potential richness of knowledge obtained through the analysis.

As for new obstacles, CA-mediated interviews will allow us to gather larger bodies of data. Manual analysis even of a small dataset can take a considerable amount of time. You need to transcribe recordings, explore and code the data, interpret the results, and present insights. In the case of large-scale studies, it becomes a practically impossible task because human resources and budgets are limited.

CAI can solve both problems by doing automatic transcription and exploratory analysis. Automatic transcription can speed up transcription via speech-to-text technology (e.g. Google Speech-to-text service) and prior work shows almost the same quality of the approach in comparison to human transcription [33].

There are the first attempts to conduct automatic analysis both from computational and contextual perspectives. Computational aspects of the problem define if we can make an automatic analysis of qualitative data such as conversation transcripts (e.g. [19, 31, 32]. Contextual aspects of the problem define what exactly should be shown to a researcher as automatic visualization: what are the researcher needs and routine practices of analysis. Prior work presents analysis of typical scenarios and decisions that designers make about data analysis (e.g. [1, 14, 20].

When we understand the needs of researchers in the interview analysis, we will be able to build systems that support such work by summarizing the content into the topics [3, 23], extracting sentiments [6] or sequences of Part-of-Speech tags, and other more advanced methods of NLP.

4 CURRENT PROGRESS AND PLANS FOR FUTURE WORK

The PhD project has been going for 3 months. It still is in the early phases of exploration what the state-of-the-art is and what are the opportunities for the research. At the moment, together with my supervisor, I submitted an overview of the use-cases of how CAs can facilitate interviewing. An overview describes the main steps of the interviewer in the procedure and how each step can be supported by a conversational AI.

The next step is to conduct a literature review to answer RQ1 and generate a set of metrics of interview quality. Once these metrics will be generated, the next step is to test them on a meaningful research question. At the moment, the idea is to test these metrics to compare experimentally how the input type (voice or text) transforms interviewees' responses.

5 ACKNOWLEDGEMENTS

I want to thank my supervisors Evangelos Niforatos and Gerd Kortuem for their support and guidance. I am also grateful to Ilya Musabirov and Xu Han for their advises regarding this submission and KInD research group (TU Delft) for their feedback during my work on the submission.

REFERENCES

[1] Antoine Barczewski, Anastasia Bezerianos, and Nadia Boukhelifa. 2020. How Domain Experts Structure Their Exploratory Data Analysis: Towards a Machine-Learned Storyline. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.

[2] Florian Bemmanna, Ramona Schoedela, Niels Van Berkel, and Daniel Buschek. 2021. Chatbots for Experience Sampling-Initial Opportunities and Challenges. In *2021 Joint ACM Conference on Intelligent User Interfaces Workshops, ACMIUI-WS 2021*. Association for Computing Machinery.

[3] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. *the Journal of machine Learning research* 3 (2003), 993–1022.

[4] Yu-Lin Chang, Yung-Ju Chang, and Chih-Ya Shen. 2019. She is in a Bad Mood Now: Leveraging Peers to Increase Data Quantity via a Chatbot-Based ESM. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*. 1–6.

[5] Paul DiMaggio, Manish Nag, and David Blei. 2013. Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of US government arts funding. *Poetics* 41, 6 (2013), 570–606. Publisher: Elsevier.

[6] Ronen Feldman. 2013. Techniques and applications for sentiment analysis. *Commun. ACM* 56, 4 (2013), 82–89. Publisher: ACM New York, NY, USA.

[7] Helen Fitt. 2018. Researching mobile practices: Participant reflection and audio-recording in Repeat Question Diaries. *Qualitative Research* 18, 6 (2018), 654–670. Publisher: SAGE Publications Sage UK: London, England.

[8] Xu Han, Michelle Zhou, Matthew J. Turner, and Tom Yeh. 2021. Designing Effective Interview Chatbots: Automatic Chatbot Profiling and Design Suggestion Generation for Chatbot Debugging. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–15.

[9] Razan Jaber and Donald McMillan. 2020. Conversational User Interfaces on Mobile Devices: Survey. In *Proceedings of the 2nd Conference on Conversational User Interfaces (CUI '20)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3405755.3406130

[10] Soomin Kim, Joonhwan Lee, and Gahgene Gweon. 2019. Comparing data from chatbot and web surveys: Effects of platform and conversational style on survey response quality. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–12.

[11] Craig Kolb, Ask Afrika, and Peter Millican. 2006. Connecting with Elizabeth: Using artificial intelligence as a data collection aid. In *Connections, MRS annual conference*. 22–24.

[12] S. Kvale. 1996. InterViews: an introduction to qualitive research interviewing: Sage. (1996).

[13] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. 2017. *Research methods in human-computer interaction*. Morgan Kaufmann.

[14] Peter Lloyd, Almila Akdag Salah, and Senthil Chandrasegaran. 2021. How Designers Talk: Constructing and Analysing a Design Thinking Data Corpus. In *ASME 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers Digital Collection.

[15] Tittaya Mairittha, Nattaya Mairittha, and Sozo Inoue. 2019. A dialogue-based annotation for activity recognition. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*. 768–773.

[16] Rashid Minhas, Camilla Elphick, and Julia Shaw. 2021. Protecting victim and witness statement: examining the effectiveness of a chatbot that uses artificial

intelligence and a cognitive interview. *AI & SOCIETY* (2021), 1–17. Publisher: Springer.

[17] Yuuki Nishiyama and Kaoru Sezaki. 2021. Experience Sampling Tool for Repetitive Skills Training in Sports using Voice User Interface. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*. 54–55.

[18] Jay F. Nunamaker, Douglas C. Derrick, Aaron C. Elkins, Judee K. Burgoon, and Mark W. Patton. 2011. Embodied conversational agent-based kiosk for automated interviewing. *Journal of Management Information Systems* 28, 1 (2011), 17–48. Publisher: Taylor & Francis.

[19] Xin Qian, Ryan A. Rossi, Fan Du, Sungchul Kim, Eunyee Koh, Sana Malik, Tak Yeon Lee, and Joel Chan. 2021. Learning to Recommend Visualizations from Data. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. 1359–1369.

[20] Vivek Rao, Euiyoung Kim, Jieun Kwon, Alice M. Agogino, and Kosa Goucher-Lambert. 2021. Framing and Tracing Human-Centered Design Teams' Method Selection: An Examination of Decision-Making Strategies. *Journal of Mechanical Design* 143, 3 (2021), 031403. Publisher: American Society of Mechanical Engineers.

[21] Melanie Revilla and Mick P. Couper. 2019. Improving the use of voice recording in a smartphone survey. *Social Science Computer Review* (2019), 0894439319888708. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

[22] Melanie Revilla, Mick P. Couper, Oriol J. Bosch, and Marc Asensio. 2020. Testing the Use of Voice Input in a Smartphone Web Survey. *Social Science Computer Review* 38, 2 (April 2020), 207–224. https://doi.org/10.1177/0894439318810715 Publisher: SAGE Publications Inc.

[23] Margaret E. Roberts, Brandon M. Stewart, and Dustin Tingley. 2019. Stm: An R package for structural topic models. *Journal of Statistical Software* 91, 1 (2019), 1–40.

[24] Kathryn Roulston. 2010. Considering quality in qualitative interviewing. *Qualitative research* 10, 2 (2010), 199–228. Publisher: Sage Publications Sage UK: London, England.

[25] Herbert J. Rubin and Irene S. Rubin. 2011. *Qualitative interviewing: The art of hearing data*. sage.

[26] Michael F. Schober, Frederick G. Conrad, Christopher Antoun, Alison W. Bowers, Andrew L. Hupp, and H. Yanna Yan. 2020. Interacting with Interviewers in Text and Voice Interviews on Smartphones. In *Interviewer Effects from a Total Survey Error Perspective*. Chapman and Hall, 179–190.

[27] Michael F. Schober, Frederick G. Conrad, Christopher Antoun, Patrick Ehlen, Stefanie Fail, Andrew L. Hupp, Michael Johnston, Lucas Vickers, H. Yanna Yan, and Chan Zhang. 2015. Precision and disclosure in text and voice interviews on smartphones. *PloS one* 10, 6 (2015), e0128337. Publisher: Public Library of Science San Francisco, CA USA.

[28] Joseph Weizenbaum. 1966. ELIZA—a computer program for the study of natural language communication between man and machine. *Commun. ACM* 9, 1 (1966), 36–45. Publisher: ACM New York, NY, USA.

[29] Ziang Xiao, Michelle X. Zhou, Wenxi Chen, Huahai Yang, and Changyan Chi. 2020. If I Hear You Correctly: Building and Evaluating Interview Chatbots with Active Listening Skills. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[30] Ziang Xiao, Michelle X. Zhou, Q. Vera Liao, Gloria Mark, Changyan Chi, Wenxi Chen, and Huahai Yang. 2020. Tell me about yourself: Using an AI-powered chatbot to conduct conversational surveys with open-ended questions. *ACM Transactions on Computer-Human Interaction (TOCHI)* 27, 3 (2020), 1–37. Publisher: ACM New York, NY, USA.

[31] Kristina Y. Yordanova, Burcu Demiray, Matthias R. Mehl, and Mike Martin. 2019. Automatic detection of everyday social behaviours and environments from verbatim transcripts of daily conversations. In *2019 IEEE International Conference on Pervasive Computing and Communications*. IEEE, 1–10.

[32] Michelle X. Zhou, Gloria Mark, Jingyi Li, and Huahai Yang. 2019. Trusting virtual agents: The effect of personality. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 9, 2-3 (2019), 1–36. Publisher: ACM New York, NY, USA.

[33] Kirsten Ziman, Andrew C. Heusser, Paxton C. Fitzpatrick, Campbell E. Field, and Jeremy R. Manning. 2018. Is automatic speech-to-text transcription ready for use in psychological experiments? *Behavior research methods* 50, 6 (2018), 2597–2605. Publisher: Springer.