Towards a holistic approach and measurement of humans interacting with speech-based technology

ASTRID CAROLUS, Media Psychology, Institute of Human-Computer-Media, University of Würzburg, Germany
CAROLIN WIENRICH, Human-Technique Systems, Institute of Human-Computer-Media, University of Würzburg, Germany

CCS Concepts: • **Human-centered computing → Natural language interfaces**; • **Computing methodologies → Philosophical/theoretical foundations of artificial intelligence**; • **Social and professional topics → User characteristics**.

Additional Key Words and Phrases: speech-based technology, user experience, social science, conceptualization, methodology

1 TECHNOLOGY AS PSYCHOLOGICAL RELEVANT INTERACTION PARTNERS

For a long time, the perspective of both public as well as scientific discussions of successful digitalization was limited to technical equipment and technical competences. If the human users were considered, it would rather be in a limited way focusing only on the ability to think rationally. Humans were predominantly referred to in the light of the "Homo oeconomicus" emphasizing the rational capabilities of humans in terms of their rational thinking, decision-making and action. Underlying emotional or motivational aspects were rather disregarded. From this perspective, technology was conceptualized as a tool that humans use to aim at their specific targets and to achieve their goals. Accordingly, interactions with technology were usually evaluated according to pragmatic aspects. The usability approach reflects this focus on the efficient and effective achievement of goals thereby meeting the requirements this "rational perspective" postulates. However, research has questioned this limited perspective on humans interacting with media and computer since the 1990s. Conceptualizing "computers as social actors" (CASA) studies have shown that technological devices trigger social reactions in their human users that are similar to those in human-human interactions. With their basic assumption that "media equals real life", Reeves and Nass (1996) introduced their concept of "media equation" postulating that people involuntarily tend to treat media entities like real people [9]. Analyzing media devices (e.g., computers, smartphones, robots) and persons appearing in media content (e.g. movie characters, news anchor), studies in this research area showed that social norms, which apply to human-human interactions and which are well-studied in social science research also apply in the context of media (see also "parasocial interaction"; [4]). Other studies revealed that social-cognitive processes can be transferred to digital interaction partners [11, 14].

2 WHAT DO I KNOW ALREADY?

This widening of perspectives on the interaction of humans and technology resulted in the conceptualization of digital technologies as psychologically relevant counterparts users do not only use but interact with. Referring to the high frequency of use in terms of recurring interactions with the device, Carolus et al. (2019) investigated a possible social relationship of the users with their smartphones, which they referred to as "digital companions" [2]. Besides the pragmatic functions the phone offers, they provided empirical evidence that the users seem to feel like they were in a social relationship with their device. Following the tradition of CASA studies, characteristics of social relationships (e.g., closeness, trust) and their outcomes (e.g., stress, coping with stress) were adapted to yield a model of human–smartphone relationships, which was empirically confirmed. As a result, and with the focus on the human users' cognitions, emotions and actions, human-technology interaction gets closer to human-human interaction. The ongoing

technological progress underlines the approximation insofar as technology itself seems to get closer to the human being. Modern technology adopts "human-like" attributes, which further strengthens the (conscious and unconscious) perception that the technological counterpart “equals real life” [3, 13].

Consequently, the former “rational perspective” on human users and their interactions with technologies are fundamentally expanded by a more holistic understanding, which involves not only the human cognitive capabilities but also the emotional and the motivational functioning as well as social needs and mechanisms. The concept of user experience (UX) does more justice to this complexity in that it more strongly "encompass[s] all emotions, perceptions, preferences, perceptions, physiological and psychological reactions, behaviors, and performances [that] occur before, during, and after use" [1, 6]. However, the majority of methodological approaches and measures in this area is limited to hedonic aspects, such as pleasure in interaction or stimulation [5, 7].

Recent contributions increasingly consider eudaimonic aspects, aiming for personal goals through technology use in terms of need fulfillment or meaning [8] or social aspects, aiming for relatedness [10, 12]. Although recent UX studies have expanded the perspective substantially, they still miss to meet the whole spectrum of requirements the early and the more recent CASA studies as well as the recent technological improvements suggest. Human-computer interaction, which increasingly resembles human-human interaction, expands its power and its sphere of influence. Just as human interactions are understood as social encounters, which affect the human on cognitive, emotional and motivational levels, human-computer interactions must also be conceptualized and analyzed in these dimensions. In this way, first studies revealing experiences that cannot be classified as either hedonic or eudaimonic but are coded as "social" experiences are regarded as first indications of this idea [6, 8]. However, a systematic reappraisal aiming for a holistic theoretical as well as methodological approach to human-technology interaction is still pending.

3 HOW DO I STUDY THE PHENOMENON?

As social scientists we conduct experimental studies in the laboratory as well as online or conduct (online) surveys. Most studies are limited to only one measurement point. Long-term studies are just as rare as indirect survey methods (e.g., implicit testing, physiological measures) or field observations. A typical approach is to manipulate the appearance of the technological counterpart, for example, to analyze the participants’ reaction to it. In an online experiment we asked for the capacity of smart speakers to elicit empathy and showed that participants watching a smart speaker being treated rudely results in significantly higher ratings of empathy with the device [3]. In another study, different human-robot interactions in a work-context were simulated in virtual reality. The results revealed strong effects of the robot design on need fulfillment and gender [12].

4 WHAT I WOULD LIKE TO KNOW?

With an emphasis on theory and methodological implications we would like to learn more about how researchers from different disciplines tackle the challenges in the area of speech-based systems. We would like to provide our holistic perspective (1) on HCI in general and (2) humans interacting with speech-based technology in particular to researchers from other disciplines to get to know their point of view.

5 WHAT DO I WANT TO LEARN FROM DIFFERENT DISCIPLINES?

As research on speech-based technology is only at the beginning of the process to become an established area of research we recognize a momentum to address the research subject in its broadness to develop research programs, which combine the expertise from the variety of disciplines involved in the development, analysis or understanding of

speech-based technology – but also the disciplines being rather indirectly affected (e.g. pedagogy, gerontology). In this sense, we want to derive and develop research concepts and approaches that are truly interdisciplinary and, long-term oriented and sustainable. One very first idea could be the collection and analysis of the theoretical foundations (and methodological approaches) in the field to provide a first overview of the status quo and to lay the foundation for future research in the field.

6 WHAT DO WE WANT TO TEACH OTHER DISCIPLINES?

We want to broaden the perspective on user experience to meet the whole spectrum of requirements the early and the more recent CASA studies as well as the recent technological improvements suggest.

REFERENCES

[1] Javier A Bargas-Avila and Kasper Hornbæk. 2011. Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 2689–2698.

[2] Astrid Carolus, Jens F Binder, Ricardo Muench, Catharina Schmidt, Florian Schneider, and Sarah L Buglass. 2019. Smartphones as digital companions: Characterizing the relationship between users and their phones. *New Media & Society* 21, 4 (2019), 914–938.

[3] Astrid Carolus, Carolin Wienrich, Anna Toerke, Tobias Friedel, Christian Schwietering, et al. 2021. 'Alexa, I feel for you!'-Observers' Empathetic Reactions towards a Conversational Agent. *Frontiers in Computer Science* 3 (2021), 46.

[4] David C Giles. 2002. Parasocial interaction: A review of the literature and a model for future research. *Media psychology* 4, 3 (2002), 279–305.

[5] Marc Hassenzahl, Michael Burmester, and Franz Koller. 2003. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In *Mensch & computer 2003*. Springer, 187–196.

[6] Marc Hassenzahl, Sarah Diefenbach, and Anja Göritz. 2010. Needs, affect, and interactive products–Facets of user experience. *Interacting with computers* 22, 5 (2010), 353–362.

[7] Mare Hassenzahl, Axel Platz, Michael Burmester, and Katrin Lehner. 2000. Hedonic and ergonomic quality aspects determine a software's appeal. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 201–208.

[8] Elisa D Mekler and Kasper Hornbæk. 2016. Momentary pleasure or lasting meaning? Distinguishing eudaimonic and hedonic user experiences. In *Proceedings of the 2016 chi conference on human factors in computing systems*. 4509–4520.

[9] Byron Reeves and Clifford Nass. 1996. *The media equation: How people treat computers, television, and new media like real people*. Cambridge university press Cambridge, United Kingdom.

[10] Carolin Wienrich and Johanna Gramlich. 2020. AppRaiseVR–An Evaluation Framework for Immersive Experiences. *i-com* 19, 2 (2020), 103–121.

[11] Carolin Wienrich, Richard Gross, Felix Kretschmer, and Gisela Müller-Plath. 2018. Developing and proving a framework for reaction time experiments in VR to objectively measure social interaction with virtual agents. In *2018 IEEE conference on virtual reality and 3D user interfaces (VR)*. IEEE, 191–198.

[12] Carolin Wienrich and Marc Erich Latoschik. 2021. eXtended Artificial Intelligence: New Prospects of Human-AI Interaction Research. *Frontiers in Virtual Reality* (2021).

[13] Carolin Wienrich, Clemens Reitelbach, and Astrid Carolus. 2021. The Trustworthiness of Voice Assistants in the Context of Healthcare: Investigating the Effect of Perceived Expertise on the Trustworthiness of Voice Assistants, Providers, Data Receivers, and Automatic Speech Recognition. *Frontiers in Computer Science* 3 (2021), 53.

[14] Carolin Wienrich, Kristina Schindler, Nina Döllinqer, Simon Kock, and Ole Traupe. 2018. Social presence and cooperation in large-scale multi-user virtual reality-the relevance of social interdependence for location-based environments. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 207–214.