The Impact of the Researcher's Social Presence on Synchronous Remote Usability Testing

Eun-Jeong Kang
Cornell University
Ithaca, USA
ek646@cornell.edu

Xin Chen
Cornell University
Ithaca, USA
xc297@cornell.edu

Xianyi Li
Cornell University
Ithaca, USA
xl652@cornell.edu

ABSTRACT

Synchronous remote usability testing is widely used by HCI researchers and UX designers to investigate users' thoughts about prototypes. Researchers can host online meetings to conduct one-on-one interviews or usability tests conveniently. Although remote usability testing offers convenience to both researchers and participants, recent research shows that participants can be influenced by the experimental mode. To understand how varied usability test settings affect participants as well as test results, we conducted a comparison study of three researcher presence modes, focusing on the researcher's social presence. In a between-subjects experiment with 32 participants, we compared three conditions – (1) Camera, (2) Avatar, and (3) No-camera. The results show that there was no significant difference in usability test results. However, we found that participants tended to ask more questions and complete tasks more quickly in the camera setting. Based on these results, we discuss which setting would be beneficial for researchers in remote usability tests and how the avatar setting could be applied to user studies as an alternative.

CCS CONCEPTS

• **Human-centered computing → User studies**.

KEYWORDS

user studies, usability test, social presence

ACM Reference Format:
Eun-Jeong Kang, Xin Chen, and Xianyi Li. 2022. The Impact of the Researcher's Social Presence on Synchronous Remote Usability Testing. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX).* ACM, New York, NY, USA, 25 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION AND BACKGROUND

As the COVID-19 pandemic has continued, UX designers and HCI researchers who conduct user-centered studies have increasingly adopted remote synchronous usability testing methods [7, 16]. Video conferencing platforms (e.g., Zoom, Google Hangouts) have become the primary way for many people to work or socialize [4], and remote usability tests are likewise conducted synchronously over these tools. In general, researchers and participants meet via an online conference tool, and participants then perform tasks under the researcher's moderation [5]. Conducting user studies over such online platforms allows both parties to meet regardless of location and lets participants schedule sessions with researchers flexibly [12].

Despite these advantages, remote usability tests are easily affected by the elements surrounding participants [21]. The underlying principles of remote usability testing are largely the same as those of laboratory usability testing [19]. However, researchers have no control over participants' varied environments. For example, participants may be distracted by which online conference tool is used or by how the researcher appears on screen. Furthermore, during meetings held over online conference tools, people attend to other participants rather than focusing on the content, and are sometimes distracted by activity outside the screen [8]. Given that usability testing aims to probe target users' behaviors with an interface or prototype in order to test the researcher's hypothesis [10], it is important to know whether such surrounding factors affect these studies.

A substantial body of research in the HCI community has examined what constitutes an ideal study setting. Alhadreti conducted a comparison study between synchronous and asynchronous remote usability tests [2]. That study is relevant to our research in that the synchronous condition was conducted primarily over an online conference tool, but it did not account for the researcher's own presentation within the conditions. Voit et al. explored five experimental methods (online, AR, VR, lab, and in-situ) to compare evaluation results using early prototypes [21]. However, that research evaluated smart artifacts rather than websites, and did so in a synchronous setting.

Specifically, the researcher's social presence may be a factor affecting participants and usability test results. Social presence is the "sense of being with another" [3]; according to social presence theory, it enables people to perceive others as salient through interaction in interpersonal relationships [17]. This concept allows people to feel connected with others, and people tend to engage more in a situation when they perceive a high level of social presence [18]. For example, students have been asked to turn on their cameras because camera use is likely to increase their engagement in the learning situation [1]. Similarly, in the case of virtual assistants, people tend to control the agent more dynamically when they perceive a higher social presence [15]. A large body of research has already examined social presence, but it has not yet been explored in the context of usability testing. We therefore ask the following question: *"Does the researcher's social presence affect synchronous*

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

remote usability testing?" Following [1], we controlled the level of researcher visibility over the camera to answer this question. Based on prior work [1, 21], we formulated three hypotheses:

(1) **H1.** Researcher's visibility affects interface usability result
(2) **H2.** Researcher's visibility affects participants' performance in completing a task
(3) **H3.** Researcher's visibility affects participants' engagement with the researcher

To test these hypotheses, we examined how the researcher's social presence influences participants in a synchronous remote usability test. We conducted a comparison study with three controlled settings for how the researcher appeared to participants: (1) camera, (2) avatar, and (3) no camera. Using the online conference tool Zoom, participants were asked to complete two tasks on a website while thinking aloud. After finishing the tasks, they evaluated the website in a post-survey. In total, 32 participants took part in our experiment.

The results show no substantial difference in website evaluations across the three conditions. However, we found that participants interacted with the researcher through questions more often and completed tasks considerably faster in the (1) camera condition. The (2) avatar condition ranked second on both measures, and the (3) no-camera condition ranked last. In summary, when the researcher disclosed their appearance to participants, tasks were completed more efficiently and interaction between researcher and participant increased. Based on this study, we discuss which settings work best for HCI researchers in remote usability tests, and consider avatar-mediated communication as an alternative to camera settings. We believe our research contributes to the development of HCI methodology and broadens the discussion of remote usability testing.

2 METHODOLOGY

We conducted a comparative study to examine how a researcher's social presence affects participants. We manipulated the researcher's social presence by controlling the researcher's camera setting. The study used a between-subjects design. In the experiment, participants were asked to evaluate an existing website by completing our tasks. Afterward, they answered a post-survey and open-ended questions. The study was conducted online, and participants interacted with the website using their own computers and internet connections. All study procedures were reviewed by the Institutional Review Board (IRB).

2.1 Materials

The remote usability test was conducted over the online conference software Zoom [1], which is widely used for meetings and educational purposes. The platform allowed participants to remotely control the researcher's screen and interact with the website, and it allowed the experimenter to observe participants' interactions with the website and to communicate with them through audio, video, and avatar. For the avatar condition, we used the animal avatars provided by the tool. These avatars mirror the user's facial expressions and head movements.

The website used for the experiment was the Frans Hals Museum website [2]. We chose this website for three reasons. First, its design differs from that of conventional websites, emphasizing artistic interactions. Second, it provides a booking system with a structured flow as well as an information section that invites exploration. Third, it has a known usability problem in the ticket-booking process: the page translation is inconsistent, such that English speakers cannot complete a booking. We therefore expected this website to elicit the range of behaviors our study was designed to observe.

2.2 Procedure

After signing up, participants accessed the online meeting link using their own laptop. Before the experiment began, the procedure was introduced to the participant and consent was obtained. The researcher then granted screen-control permission to the participant so that they could carry out the tasks.

Participants were asked to complete two tasks on the website:

- Imagine you would like to visit the Frans Hals Museum with your younger brother or sister. You have heard that the museum has a painting workshop. Could you 1) find more information about the workshop on the museum website, and 2) reserve a ticket for the event?
- Frans Hals is a famous artist, but you are not very familiar with his artwork. Could you find an introduction to Frans Hals and the timeline of his life?

During the tasks, participants in the avatar group were shown an avatar representing the experimenter, participants in the camera-on group were shown a live video of the experimenter, and participants in the camera-off group were shown no video. The researcher asked participants to interact with the website while continuously thinking aloud and verbalizing their thoughts as they moved through the user interface. All on-screen behavior was recorded.

After completing the tasks, participants were asked to complete a post-survey (Section 2.4) through Qualtrics, along with open-ended questions. In the open-ended questions, we asked about their overall experience and about the researcher's social presence during the session.

2.2.1 Researcher's behavior. Because our study focused on researcher visibility, it was important to standardize how the researcher behaved toward participants. While participants performed the tasks, the researcher did not prompt them unless they asked for help or posed a question. This was done because the researcher's behavior could otherwise have influenced participants.

2.3 Participants

Participants were recruited through the institutional social science research pool (SONA). All participants were pre-screened, were over the age of 18, and were students at Cornell University. We successfully recruited 32 participants (n = 32; 11 male and 21 female). After sign-up, we randomly assigned them to the three conditions. Detailed demographic information is presented in Table 1.

[1] https://zoom.us/

[2] https://www.franshalsmuseum.nl/en/

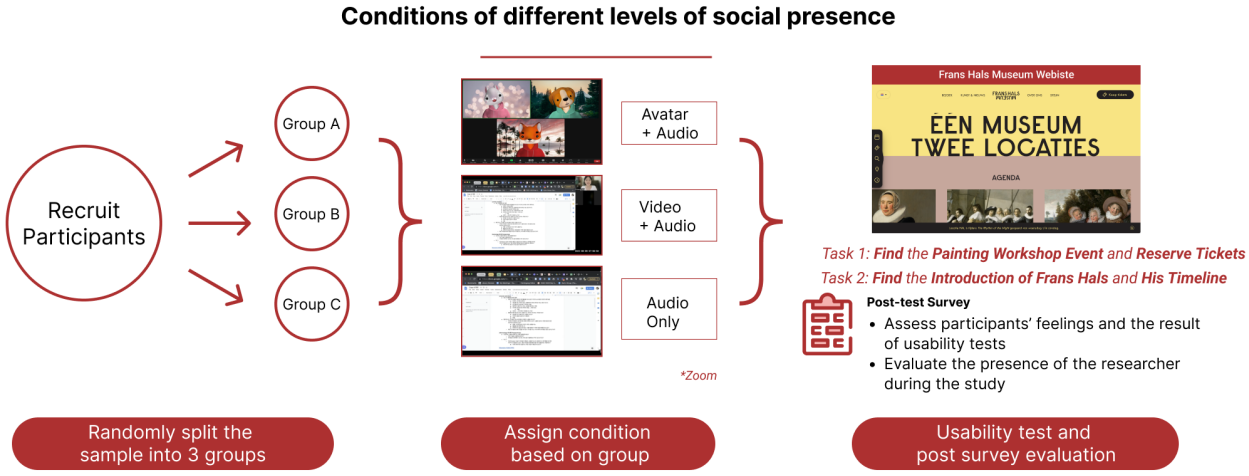

Figure 1: Overall process of our experiment. After recruitment, participants were assigned to one of three conditions: (1) camera, (2) avatar, or (3) no-camera. After finishing the experiment, participants answered a survey and open-ended questions.

	Male	Female
Camera	3	9
Avatar	5	5
No-Camera (Audio Only)	3	7

Table 1: Participants Demographics and Condition Assignment

2.4 Measurement

Following [1], we controlled the level of researcher visibility over the camera. We measured one outcome for each of our three hypotheses:

(1) **H1.** Researcher's visibility affects interface usability result
(2) **H2.** Researcher's visibility affects participants' performance in completing a task
(3) **H3.** Researcher's visibility affects participants' engagement with the researcher

2.4.1 System Usability Scale (SUS) [H1]. We used the System Usability Scale (SUS) [6] in our post-survey to measure perceived interface usability (H1). The SUS consists of a 10-item questionnaire; an interface scoring above 70 is generally considered to have 'good' usability. It is widely used by HCI researchers and UX designers to evaluate a broad range of products and services. We administered the SUS items on a 5-point Likert scale (strongly disagree – strongly agree).

2.4.2 Task completion time [H2]. We measured task completion time for each task to assess participant performance (H2), following Voit et al. [21]. Using the timestamps recorded in the transcripts, we manually measured each task's completion time: specifically, the interval between the researcher announcing the start of the task and the participant finishing it.

2.4.3 Questions asked by participants [H3]. We counted the number of questions participants asked while performing each task in order to examine their interaction with the researcher (H3). We counted one question each time a participant addressed a question to the researcher during a task.

2.5 Data analysis

2.5.1 Quantitative analysis. We analyzed all of the measures described in Section 2.4 using descriptive statistics (mean, SD), one-way ANOVA, and pairwise t-tests.

2.5.2 Qualitative analysis. We conducted a qualitative analysis of the open-ended responses. We focused on examining (1) whether the researcher's setting influenced participants during the usability test, and (2) which setting participants would prefer. Working from transcripts of the audio recordings, which we anonymized beforehand, each author independently open-coded the responses along these two themes. After initial coding and discussion, we merged and organized the codes.

2.6 Ethical considerations

The study procedure was reviewed and approved by the Institutional Review Board (IRB). All participants provided informed consent prior to participating, and their personal information was kept confidential.

3 RESULTS

We analyzed three measures to investigate our hypotheses: (1) a comparison of System Usability Scale (SUS) scores across the

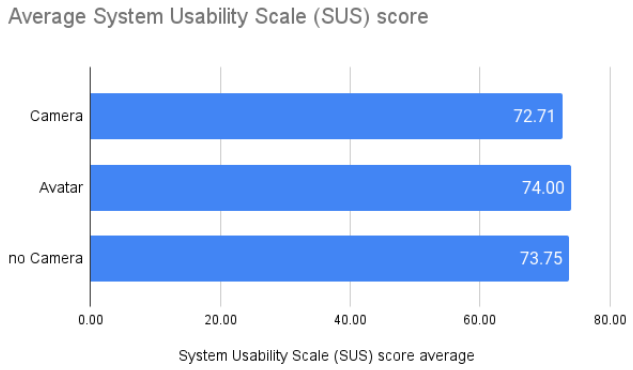

Figure 2: Average System Usability Scale (SUS) score by condition

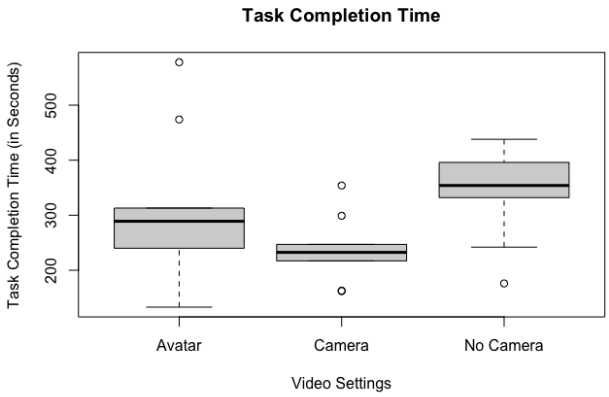

Figure 3: Task completion time by condition

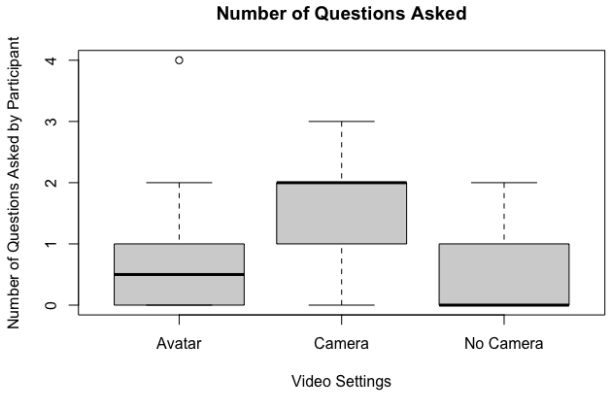

Figure 4: Number of questions asked by participants, by condition

three conditions, (2) a comparison of task completion time across the three conditions, and (3) a comparison of participants' level of engagement across the three conditions. We then report the qualitative themes shared in the open-ended responses.

3.1 System Usability Scale (SUS)

As seen in Figure 2, the average score for the avatar condition was 74 (SD = 6.687). For the camera condition, the average score was 72 (SD = 6.696). The average score for the no-camera condition was 73.75 (SD = 6.263). We conducted an ANOVA to determine whether there were significant differences among the means of the three conditions. The null hypothesis of no difference among the means could not be rejected, indicating that there were no statistically significant differences among the three conditions. This suggests that the researcher's camera setting did not significantly affect participants' perception of the web design (H1).

3.2 Task Completion Time

As seen in Figure 3, participants in the condition where the researcher turned on the camera completed the usability tasks faster. For task 1, the average completion time was 3:18 (SD = 1:07) in the camera condition, 3:40 (SD = 1:50) in the avatar condition, and 4:24 (SD = 1:19) in the no-camera condition. For task 2, the average completion time was 0:56 (SD = 0:25) in the camera condition, 1:23 (SD = 0:37) in the avatar condition, and 1:11 (SD = 0:37) in the no-camera condition. We conducted t-tests to compare task completion time between conditions. The results showed a statistically significant difference between the camera and no-camera conditions: participants completed tasks significantly faster when the researcher's camera was turned on (p = 0.0269). These results suggest that the researcher's visible presence may affect participants' motivation and focus, leading to faster task completion.

3.3 Questions by Participants

As seen in Figure 4, participants were more likely to ask the researcher questions when the researcher was visible on camera. The average number of questions asked during the tasks was 1.583 in the camera condition, 0.9 in the avatar condition, and 0.7 in the no-camera condition. The corresponding totals were 19, 9, and 7 questions. Pairwise t-tests revealed a statistically significant difference in the number of questions asked between the camera and no-camera conditions, with participants asking significantly more questions when the researcher's camera was turned on (p = 0.0267).

3.4 Summary of Quantitative Analysis

In summary, the researcher's visibility had no measurable impact on usability test scores (H1), but it did affect task efficiency (H2) and participant engagement (H3).

3.5 Qualitative Analysis

3.5.1 No perceived effect on results. Most participants believed that the researcher's presence did not interfere with their behavior; they tended to concentrate on completing the tasks rather than attending to the researcher. One participant stated: *"I was trying to complete the task as the researchers asked me, so I am confident that I finished the tasks without being affected by researcher's setting."* (P13 – no camera) Even in the no-camera setting, because it is common for

meeting partners to turn their cameras off, participants understood the researcher's situation and tried to stay focused on the study.

3.5.2 Confidence in User studies. Although participants did not believe the camera setting affected their results, most of them (N=26) preferred the experimental setting with the camera.

> "I would just say turning on camera is probably the most ideal. Because it'll be helpful to answer the question." (P17 – camera)

When the researcher turned off the camera, participants felt as though they were talking to themselves during the test, which left them feeling disconnected. Participants also wanted supervision and affirmation while working, because they hoped to complete the tasks accurately. When the researcher's camera was on, participants felt that the researcher would guide them if they became lost.

In the avatar condition, some participants said they were **surprised**, as researchers do not typically use avatars. Notably, participants who themselves preferred not to turn on a camera favored the avatar over the camera setting, because they did not want to feel surveilled by the researcher.

> "Might feel pressure if (researcher) camera is on, I (would feel like) someone is seeing me." (P4 – avatar)

Conversely, a participant who preferred cameras to be on also viewed the avatar setting positively, but for a different reason: they felt the avatar mitigated the sense of disconnection better than turning the camera off.

> "Cat visualization is better than a black screen. Similar to actual faces is comforting in a way" (P25 – avatar)

However, participants found it hard to read the researcher's expression through the avatar. The avatars were also imperfect and could not fully reproduce the researcher's movements. As a result, participants reported that they sometimes over-interpreted the researcher's expression or tried to visualize the researcher based on their voice alone.

4 DISCUSSION

In what follows, we discuss our results in relation to our research question and consider which settings benefit researchers conducting user studies remotely. We also propose avatar-based remote usability testing as an alternative approach for researchers.

4.1 Suggestions for Researchers' Settings in Usability Tests

Regarding H1, we found that usability test results are not substantially affected by the researcher's visibility. Voit et al. similarly found that the research method does not affect the assessment of prototype usability [21]. Regardless of how the researcher presented themselves, participants evaluated the system comparably and identified the same usability problems. Based on our qualitative analysis, this may be because participants focused primarily on completing the tasks.

On the other hand, depending on the type of user study, the researcher's camera setting and other factors influencing social presence merit consideration (H2, H3). Our finding that participants interacted more with the researcher in the camera condition aligns with research on social presence in other contexts [15, 20]. The degree of social presence affects learners' levels of interaction in online learning [20], and Shao and Kwan found that the social presence of a voice assistant leads users to control the agent more dynamically [15]. Questions and dynamic interaction matter in usability testing, depending on the researcher's goals [10]. If a study involves think-aloud or interview components in which participants must be encouraged to share their thoughts, a 'camera-on' setting is more likely to open them up than a setting without a camera – much as researchers commonly use ice-breaker questions to help participants disclose their thoughts.

Taken together, usability test results are not affected by the researcher's appearance, but we encourage researchers to give participants a stronger sense that the researcher is present with them, depending on the purpose of the usability test.

4.2 Expression-captured Avatar as an alternative approach for organizational communication

Expression-captured avatars have been used to represent users for various purposes in computer-mediated communication, such as messaging (Memoji in iOS) [13] and live streaming (Virtual YouTubers) [11]. Although avatars have long been used in games and chat, expression-captured avatars additionally reproduce users' physical expressions, including facial expressions and movements [13]. For this reason, they convey greater social presence than a setting without a camera [9]. In our study, we found that such expression-captured avatars could potentially serve as a medium for usability tests, accommodating both participants who feel pressured by being observed and those who feel disconnected without the researcher's visible presence.

Avatar-mediated communication may also suit researchers who wish to protect their own privacy from participants. In Lu et al.'s work, live streamers who use virtual characters feel comfortable disclosing experiences because they do not have to show themselves [11]. Likewise, junior researchers might feel less pressure when beginning a study if they use an avatar. Researchers could also select an appearance tailored to their participants. For example, a researcher might choose a well-known character when running user studies with young children in order to increase engagement. Using AI-generated characters, Pataranutaporn et al. demonstrated that such portrayals can be used productively for health checkups with older adults and for education with young children [14]. In this study we focused only on the animal avatars provided by an online conference tool; future research could explore how avatar appearance might be used to increase participant engagement.

However, as our qualitative analysis shows, avatar-mediated communication has limitations in expressing emotion. Avatars may fail to deliver social cues accurately because they render expression abstractly, and people may misinterpret them. The avatars we used displayed only simple head movements. Park and Suk demonstrated that Memojis convey emotion insufficiently through facial expression alone [13]. Similarly, if researchers adopt avatars as part of a research method, participants may not grasp the researcher's intent during the experiment. Future studies could investigate how to

improve avatar-mediated communication in the context of usability testing.

5 LIMITATIONS

We acknowledge several limitations in our study and believe that follow-up research can address them. First, we collected sparse data, which limited our ability to assess statistical significance. Given the nature of the experiment and our limited time, we used a between-subjects design and were unable to recruit as many participants as such a design warrants. A larger sample would support more robust statistical analysis.

Second, we used a single online conference tool (Zoom) because it is widely adopted, and we recruited through an institutional research participant system (SONA). Our participants were undergraduates majoring in Information Science or Communication. The results may therefore not generalize to other tools or populations.

Third, the avatar used in our study was Zoom's default feature. This feature is still in beta and is not fully reliable: unexpected problems can occur, such as the avatar suddenly disappearing. This may have affected our results to some extent.

Finally, the tasks were subjectively easy for our participants. Because most were undergraduates studying Information Science, they completed the usability tasks with little difficulty even though the website is hard to navigate.

Future research in this area should adjust the study design in several ways. First, the usability tasks should be more difficult; a well-calibrated task should involve more steps and should not be trivially completable. Second, where possible, researchers should design their own web prototypes for usability testing. Doing so removes the possibility that participants are already familiar with the site and eliminates a confounding variable. In future work, we hope to recruit a larger and more diverse participant pool and to run more in-depth experiments with a more complete experimental setup, which would allow us to examine the significance of these results more effectively and to identify factors overlooked in the present study.

6 CONCLUSION

In this study, we investigated how a researcher's social presence affects remote usability tests. The results showed no significant difference in usability test scores across the three conditions (camera, avatar, and no-camera). However, participants asked more questions and completed tasks more quickly in the camera setting, indicating that camera use may support more effective communication during remote usability testing.

These findings have implications for researchers conducting remote usability testing, suggesting that camera use may help establish an effective communication channel with participants. At the same time, the avatar setting may serve as a useful alternative, and researchers could select avatars matched to the interests of a particular participant group.

Overall, the results of this study highlight the importance of considering the impact of communication modes on the user experience in remote usability testing, and suggest that researchers should choose the most appropriate mode based on the goals and objectives of their study.

REFERENCES

[1] Holiday E Adair, Jeanne M Slattery, and Holiday Adair. 2021. Cameras in Zoom: Should They Be On. *Hand in Hand* (2021).

[2] Obead Alhadreti. 2022. A Comparison of Synchronous and Asynchronous Remote Usability Testing Methods. *International Journal of Human–Computer Interaction* 38, 3 (2022), 289–297.

[3] Frank Biocca, Chad Harms, and Judee K Burgoon. 2003. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators & virtual environments* 12, 5 (2003), 456–480.

[4] Vinton G. Cerf. 2020. Implications of the COVID-19 Pandemic. *Commun. ACM* 63, 6 (may 2020), 7. https://doi.org/10.1145/3397262

[5] Kapil Chalil Madathil and Joel S Greenstein. 2011. Synchronous remote usability testing: a new approach facilitated by virtual worlds. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2225–2234.

[6] Mandy R Drew, Brooke Falcone, and Wendy L Baccus. 2018. What does the system usability scale (SUS) measure?. In *International Conference of Design, User Experience, and Usability*. Springer, 356–366.

[7] Mark Embrett, Rebecca H Liu, Katie Aubrecht, Andriy Koval, and Jonathan Lai. 2021. Thinking together, working apart: leveraging a community of practice to facilitate productive and meaningful remote collaboration. *International Journal of Health Policy and Management* 10, 9 (2021), 528–533.

[8] Joey George, Akmal Mirsadikov, Misty Nabors, and Kent Marett. 2022. What do Users Actually Look at During 'Zoom'Meetings? Discovery Research on Attention, Gender and Distraction Effects. In *Proceedings of the 55th Hawaii International Conference on System Sciences*.

[9] Devesh Gulhane. 2022. *What Face Would You Like To Have?: The Effects of an Avatar's Facial Features on Social Presence.* Master's thesis. University of Twente.

[10] Jonathan Lazar. 2017. *Research Methods in Human Computer Interaction* (2nd edition ed.). Elsevier.

[11] Zhicong Lu, Chenxinran Shen, Jiannan Li, Hong Shen, and Daniel Wigdor. 2021. More kawaii than a real-person live streamer: understanding how the otaku community engages with and perceives virtual YouTubers. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.

[12] Edward McFadden, Danny R Hager, Claude J Elie, and Jason M Blackwell. 2002. Remote usability evaluation: Overview and case studies. *International journal of human-computer interaction* 14, 3-4 (2002), 489–502.

[13] Minjung Park and Hyeon-Jeong Suk. 2022-09-28. New Mobile Adaptation System for Better Avatar-mediated Communication; Facial Expressions in Memoji. In *Adjunct Publication of the 24th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Vancouver BC Canada). ACM, 1–5. https://doi.org/10.1145/3528575.3551437

[14] Pat Pataranutaporn, Valdemar Danry, Joanne Leong, Parinya Punpongsanon, Dan Novy, Pattie Maes, and Misha Sra. 2021. AI-generated characters for supporting personalized learning and well-being. *Nature Machine Intelligence* 3, 12 (2021), 1013–1022.

[15] Chun Shao and K Hazel Kwon. 2021. Hello Alexa! Exploring effects of motivational factors and social presence on satisfaction with artificial intelligence-enabled gadgets. *Human Behavior and Emerging Technologies* 3, 5 (2021), 978–988.

[16] LeeAnne B Sherwin, Justina Yevu-Johnson, Michelle Matteson-Kome, Mathew Bechtold, and Blaine Reeder. 2022. Remote Usability Testing to Facilitate the Continuation of Research. *Studies in health technology and informatics* 290 (2022), 424–427.

[17] John Short, Ederyn Williams, and Bruce Christie. 1976. *The social psychology of telecommunications.* Toronto; London; New York: Wiley.

[18] Zaidatun Tasir and Yahya Al-Dheleai. 2019. Web 2.0 for fostering students' social presence in online learning-based interaction. *JOTSE: Journal of Technology and Science Education* 9, 1 (2019), 13–19.

[19] Katherine E Thompson, Evelyn P Rozanski, and Anne R Haake. 2004. Here, there, anywhere: Remote usability testing that works. In *Proceedings of the 5th conference on Information technology education*. 132–137.

[20] Chih-Hsiung Tu. 2002. The measurement of social presence in an online learning environment. *International Journal on E-learning* 1, 2 (2002), 34–45.

[21] Alexandra Voit, Sven Mayer, Valentin Schwind, and Niels Henze. 2019. Online, VR, AR, Lab, and In-Situ: comparison of research methods to evaluate smart artifacts. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–12.