

Using audio-mixing software to facilitate remote data collection of conversational interactions

Luke Wilsdon
University of Wolverhampton
University of Wolverhampton, UK
l.a.wilsdon@wlv.ac.uk

Maria Uther
University of Wolverhampton
University of Wolverhampton, UK
m.uther@wlv.ac.uk

Darren Chadwick
University of Wolverhampton
University of Wolverhampton, UK
d.chadwick@wlv.ac.uk

Chris Fullwood
University of Wolverhampton
University of Wolverhampton, UK
c.fullwood@wlv.ac.uk

This paper discusses the issues with using open source, audio-mixing technology to facilitate remote data collection of speech samples, which was needed specifically during the COVID-19 pandemic, but will also be beneficial in other contexts, e.g., to save on time and costs associated with travelling. We discuss practical constraints associated with remote data collection using this technology. We also consider issues around ethics, security, and data quality in using technology to record conversational interactions. We provide the example of using common tools such as MS Teams and smartphones and two types of software to conduct interviews to collect speech data as a proof-of-concept and offer further directions for research.

Keywords: Cyberpsychology, Method-adaptation, Speech data collection, Research Methods, Ethics

1. NEED FOR REMOTE DATA COLLECTION

The unprecedented COVID-19 pandemic has created exponential change both to our immediate short-term futures, and the way the world is predicted to operate in coming years. The “new normal” is a phrase that has been uttered across the reaches of cyberspace, typically referring to the state in which we will settle following this and indeed previous crises. While the most pervasive changes will likely be to that of remote-working and remote-learning, there has been a major emphasis on the need to adapt traditional research strategies to perform in exclusively cyberspaces as a result of shift in cultural working practices. According to the Office of National Statistics, 46.6% of employed individuals in the UK practised homeworking in 2020, 86% of these individuals did so as a direct result of the COVID-19 restrictions (ONS, 2020). Regarding academia, it is reported that 78% of UK Higher Education staff wish to continue with hybrid-

working (Taylor et al., 2021), higher education managers’ attitudes regarding staff homeworking are also significantly more positive now than pre-pandemic (Forbes et al., 2020). As a result, there is the need to consider practical constraints for remote data collection, as there are several aspects which need to be considered as we rely more on online technology platforms for data collection and there is a growing body of research in this field that have considered the various challenges involved (e.g., Zhang et al., 2021; Poorjam et al., 2019; Gregory et al., 2022). The following section aims to consider some key practical areas which might in turn serve as a framework to evaluate the practical considerations for optimal data collection.

2. SPEECH AND/OR VIDEO QUALITY

One area of concern that is particularly of importance to those doing research on conversational interactions is the matter of speech and/or video quality (depending on what variables are being studied). Typically, for conversational interactions where a researcher is going to analyse speech or code facial expressions or bodily gestures research would take place in a researcher-controlled environment where sound levels, recording equipment and the general physical parameters can be tightly controlled. By contrast, remote data collection allows a vast range of variability in terms of hardware, software and physical conditions (proximity to extraneous noise) that means that the control of data quality for analysis will be challenged. For the researcher conducting remote data collection, this variability can be helped by identifying and testing common tools that are likely to be used in data collection and determining which is the optimum method of data collection that reduces systemic variability. The key concept is the reduction of systemic noise. Given that as researchers, we cannot fully eliminate extraneous noise, the goal is to reduce the foreseeable, systemic noise that is most likely to affect data quality. Recent research in this area (Zhang et al., 2021) showed that lossless formats from phones resulted in higher quality compared to video conference (Zoom) recordings, but further research is needed to determine whether these differences might result in significant variation that would affect the measurement of key variables. One would expect that the extent of the impact would depend on which variable was being measured (e.g., formant frequency, intensity, etc.). Of course, this is just one aspect, and there are other considerations (e.g., ethics, data storage and security) which are discussed below.

3. ETHICAL, DATA STORAGE/SECURITY ISSUES

Many universities and other professional bodies (e.g., British Psychological Society) altered research guidance provision within the context of COVID-19 to ensure research could still be conducted in a safe and compliant way (FEHW Ethics Resources - University of Wolverhampton, 2022; Chenneville & Schwartz-Mette, 2020). Qualitative researchers specifically may have been forced to rethink their research approach a lot more drastically than their quantitative counterparts. Qualitative research faced significantly more restrictions on both methods, and accessing their typical participant base – e.g., see Howlett (2021).

We will not labour existing ethical best practice in research (e.g., using American Psychological Association, British Psychological Society, or other similar professional bodies). Instead, we wish to highlight some specific considerations around ethical issues that are likely to crop up with collection of conversational interactions conducted remotely. The first aspect concerns confidentiality and retaining anonymity. In many forms of data collection (e.g., survey data), it is an arguably straightforward matter to keep data separate from identity. However, with speech or video data, the identity of the individuals can be compromised simply by a person recognising a particular voice or face. In terms of best practice, raw audio and/or video should only be kept on secure storage accessible to named researchers. Furthermore, unless required for quality control purposes, the ideal is that raw and unanonymised data be destroyed following preparation for analysis, it is common and often necessary for researchers to retain data for an extended period of time after publication dependent on institutional policy (5-10 years). This data retention allows reanalysis and republication at a later date and, where necessary, the ability to deal with any potential ethical complaints that may arise. The platforms in which this data is collected should also be end-to-end encrypted to prevent interception to unintended recipients, it is worth drawing attention to Zoom's recently published "Security Bulletin", in which it states in November of 2021 a buffer overflow vulnerability was discovered in certain products, potentially allowing a malicious actor to "crash the service, or leverage to execute arbitrary code" (Zoom, 2022a). Couple this with the discovery that Zoom was using "fake end-to-end encryption" and is currently in a class-action lawsuit means that researchers need to be extremely thorough with regards to utilising third-party services in regard to reputation, permissions and controversies. Given the recent proliferation and the alarming increase in sophistication of deep-fake video technology, it is imperative that researchers are thorough and have stringent policies and practises to ensure video-data does not be transmitted to unintended recipients. Deep-fake technology can be used for nefarious purposes and the impact is considerable to the individual, the individual's network, and the source of the breach, e.g., creating non-consensual pornography (Agarwal et al., 2020).

With respect to privacy concerns, it is also important to note that there are foreseeable privacy breaches that could compromise security (e.g., someone walking into a room unexpectedly at a critical moment in an interpersonal interview which is conducted in the person's home), To an extent, these are avoidable (e.g., priming participants about the likelihood of interruptions at home and what could be done). However, there are additional

considerations which are not always obvious to participants or researchers. For example, some applications on smartphones or PCs may grant access data stored on one's personal device that are not connected to research and which may include personal data such as contacts, purchase history, or location, as well as access to camera and microphones. Whilst these permissions might be sought by the app via the terms and conditions agreements which users may click through, these are not always understood nor even read by the users (see Gelinis et al., 2021). When using third-party applications to facilitate a research process, it is important for researchers to investigate and be aware of permissions sought from said third-party applications; and subsequently detail these to participants and ethics committees.

4. EXAMPLE SCENARIO OF SPEECH COMMUNICATION DATA COLLECTION

The continuation of interview-based research during the pandemic and presently shows that, while best practices and advice for researchers are important, they are not core to the ability to conduct this research. The proliferation of video-conferencing software, i.e., Microsoft Teams and Zoom, offers unique opportunities for data-collection and research, initially born of necessity but now is being explored for its convenience, accessibility, cost-effectiveness, and portability (Gray et al., 2020). A current issue that seems underappreciated regarding video-conferencing software is the reliance on cloud storage facilities. Of course, there is the ability to store recordings locally which is Zoom's default policy, however, this comes with its own set of issues: storage size, accessibility, lack of auto-transcription or more intricate technical features (Zoom 2022b). With respect to Microsoft Office (Stream), unless specific global/group sharing policies have been applied by IT managers and verified via a test recording via researchers, there is a risk that sensitive recordings transmitted and stored on the cloud could be viewed by someone who is unauthorised to do so, as the default setting is to allow recordings to be viewable by the whole organisation (Microsoft, 2022a).

5. OPTIMISING TECHNOLOGY TO SUIT RESEARCH NEEDS

The author developed a method in response to the pandemic that allowed the recording of system audio independently from video, allowing the recording of audio from interviews without reliance on Third-party systems cloud-based recording and storage system, i.e., Microsoft Teams, Zoom. The remote conversational data collection method was achieved (see Figure 1) by utilising two pieces of

open-source software in tandem, namely VoiceMeeter (2021) and Audacity (2021), a technological make-shift temporary adaptation. VoiceMeeter is a virtual mixing console and is used to create a "virtual speaker input" comprising of produced audio on a system, be that through an application and/or a microphone. Audacity is an audio mixer, used to record the virtual speaker input and microphone input onto one "track."

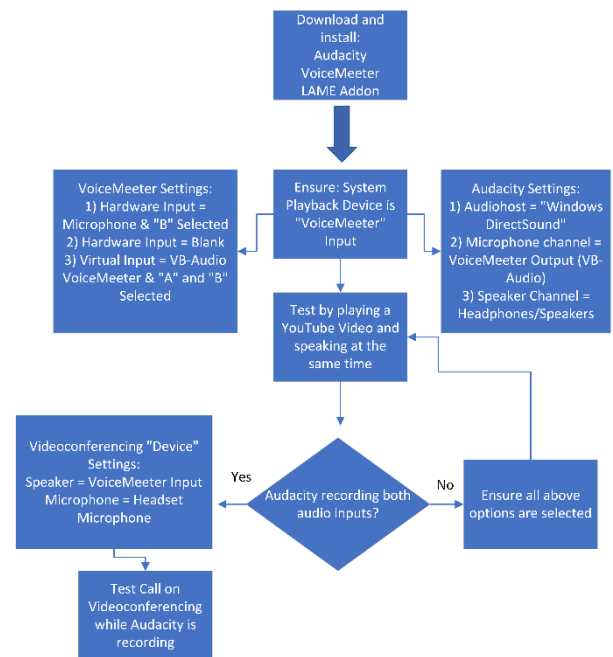


Figure 1: A setup for remote conversational data collection that reduces risk to privacy

Both MS Teams and Zoom recording systems capture both video and audio which is stored, in some capacity, on cloud storage, dependent on usage and policies. Zoom by default sets to record and store meetings locally, MS Teams' default is to store via the Cloud. The method was used for an interview study recently conducted within the University of Wolverhampton for data collection of potentially sensitive topics to bolster anonymity as well as conduct the interview remotely (Nicklin et al., 2022). The research team were tasked with weighing up the pros and cons of essentially combining two pieces of software to enable the same feature that is inherent within Microsoft Teams, Zoom or other video-conferencing providers gaining traction. Given the sensitive nature of the research, the team deemed it appropriate to eliminate any risk of unauthorised viewing born from improper organisational policies and went forward with utilising VoiceMeeter and Audacity in conjunction to allow recording of both system audio (the participant) and microphone audio (the researcher). This combination of system and microphone audio onto a single audio track allowed

a cohesive conversation to be captured and simplified the preparation for analysis on only one file, without needing to align the two separate tracks with timestamps. Also, data was prepared and analysed as a transcript using thematic analysis. Also, as video-data was unnecessary for the purposes of analysis, we deemed it appropriate to eliminate any erroneous and irrelevant data (video). It is worth noting that this solution was identified as a contextually relevant fix for a temporary issue identified within the organisation and is not the exclusive solution for remote data collection involving interviews. Open-source solutions such as OBS studio could be used to also locally record without reliance on cloud servers, though a video file is always produced, and isolation of audio would require another application/software. It must be conceded that there may be more elegant solutions to this very contextual issue, and indeed there may be issues or vulnerabilities not initially considered by the researchers. Below is a list of potential benefits identified in discussion with co-researchers.

- Suppression of the Hawthorne Effect (the idea that behaviour is modified upon recognition of being observed or attempts to “improve” their behaviour and impression solely because they are the subject of a research experiment. (Merrett, 2006)
- (a) The method allows pseudo-invisible recording, as in the visual marker of recording (red circle) is not present as there is no use of the video-conferencing recording features. It is possible that the participant may be more open to discussing sensitive topics without constant awareness of being recorded.
- Subject reactivity diminishment – subject reactivity describes the effect of a study’s procedures on its participants or subjects (Clifford & Maisto, 2000)
- (b) Onus is on the researcher – participant merely needs to be available at a specific time, no financial cost, no proximal movement, no organising babysitters, or missing work, the participant may also choose their preferred platform. Within the current context of COVID-19 and video-conferencing software ushering in a new “virtual togetherness” (Hacker et al., 2020), participating in an interview may feel like talking to a friend.

- Naturalistically and invisibly record conversational interactions in online environments, i.e., Virtual Reality Environments/Games/Social Spaces.
- (c) If video is required for VR research, an open-source solution such as OBS studio could be utilised.
- Cost-effective – this method of recording interviews requires no high-quality external microphones, merely a stable internet connection, time, and a computer/laptop with at least 4GB Ram (Table 1).
- Relative ease of use after setup and familiarisation/tests – Figure 1 above displays a flowchart guide to replicate this method

Table 1: Recommended System Requirements of each Software used in the current method

Recommended Hardware Requirements	Software			
	Microsoft Teams	Zoom	Audacity	VoiceMeeter
RAM (Memory)	4.0 GB	4.0 GB	4.0 GB	512 MB
Hard Disk (Space)	3.0 GB	Not given	Not given	Not given
CPU (Processor)	Dual-Core 1.1 GHz	Dual-Core 2 GHz	Dual-Core 2 GHz	Dual-Core 1.8 GHz

Taken from Application Websites (Microsoft, 2022b; Zoom, 2022c; Audacity 2022; VB-Audio, 2022)

6. AN EXAMPLE OF ISSUES WITHIN TECHNOLOGY-RESEARCH HYBRIDIZATION

AI Transcription is convenient and a great timesaver, but if it is stipulated in an ethics application that data will not be transmitted to third-parties or anyone who is not a named researcher, using an AI transcription service is potentially in breach of ethics agreements. Auto-transcription services relies on data to train, and subsequently test the model being implemented/developed. It may be in the terms and conditions of these services that they retain data for training and testing of models, usually with a justification that data is used for: “improvement of our products/services” – e.g., see Happyscribe’s Privacy Policy (2020). The data transmitted to these services does improve the overall effectiveness of the model, however, besides a terms and conditions agreement that data will be anonymised before testing or training, there is little users can do to verify the security of company’s respective data retention policies first-hand. Of course, companies are bound by confidentiality and GDPR agreements, however, 2021 saw the highest number of recorded data breaches (1862) with a

combined total of over 293 million victims, meaning that compliance with legislature is not absolute (ITRC, 2021) so taking extra precautions to protect one's own compliance and participant data is not a redundant measure. Couple this with the recent alert from the National Cyber Security Centre describing how the education sector is being subjected to a huge increase in ransomware attacks, the repercussions are serious and present, financially, and ethically (National Cyber Security Centre, 2021). The consequences of an interview recording being transmitted to a third-party company that is then breached is substantial - contextual depending on the sensitivity of the data collected and legibility of the data storage given a company's policy but nevertheless an ethical failing.

Audacity, the software, was acquired in May 2021 by MuseGroup and soon controversy followed. MuseGroup planned to introduce Telemetry within Audacity. Telemetry refers to a system used to gather data on the use and performance of applications to send back to the developer or allow the user to view which aids in identifying bugs and errors within software. These plans were swiftly dropped following controversy regarding MuseGroup's intention to collect IP addresses and "*data necessary for law enforcement*," and instead introduced a user-controlled error reporter with the option to send back to the developers. Further controversy followed a privacy policy change was proposed and unclear wording within the policy led some users to brand Audacity as possible "spyware." MuseGroup subsequently clarified and apologised, however, mistrust was sown, and the current researcher decided to roll back to a previous version of Audacity for the research. Another potential issue specifically related to this method is that of continuous consent, in that participants must remain consistently aware they are being recorded. The elimination of the visual indicator of recording means that participants may very well forget they are being recorded, and then it is the responsibility of the researcher to reinform and ensure participants are aware that recording is taking place. Inform clearly at the beginning of the interview and remind at set intervals that the interview is being recorded.

7. FUTURE DIRECTIONS

The internet is a tool, one in which we have been acquainted with for some time, but similarly to a mechanical tool, it evolves and shifts as context dictates. The world could not afford to be overtly sceptical of technology as it is unclear how the world would have coped without the internet in the face of the COVID-19 pandemic. It is clear, however, that the necessity born from the pandemic – for connections, service access, research, education, gaming should assist in the ushering in of a paradigm shift of research

methods. Moving away from highly expensive, regimented in-person research, to a more careful collaboration between scientists and leveraging technology for purposes above and beyond listed within this paper. Here we detail a process and experience of technology facilitation to remotely collect conversational/interview data during a time where typical face-to-face collection was impossible, we touch on the ethical issues and implications considered during the research phase, as well as how the method discussed addressed and ameliorated these. Further work is still needed to determine the effects of using different devices collect video and audio samples remotely, particularly in studies which look at specific acoustic variables or observational coding of faces and gestures. Depending on the focus of the study, different devices may affect the results in different ways.

8. REFERENCES

- Audacity Team. 2021. Audacity. (3.01). [Software]. [Accessed 18 March 2022]
- Audacity Team. (2022). Windows | Audacity®. Available at: <https://www.audacityteam.org/download/windows>
- Chenneville, T. and Schwartz-Mette, R., 2020. Ethical considerations for psychologists in the time of COVID-19. *American Psychologist*, 75(5), p.644.
- Clifford, P.R. and Maisto, S.A., 2000. Subject reactivity effects and alcohol treatment outcome research. *Journal of Studies on Alcohol*, 61(6), pp.787-793.
- Forbes, S., Birkett, H., Evans, L., Chung, H. and Whiteman, J., 2020. Managing employees during the COVID-19 pandemic: Flexible working and the future of work.
- Gelinas, L., Morrell, W., White, S.A. and Bierer, B.E., 2021. Navigating the ethics of remote research data collection. *Clinical Trials*, 18(5), pp.606-614.
- Glaser, B.G. and Strauss, A.L., 2017. *The discovery of grounded theory: Strategies for qualitative research*. Routledge.
- Gregory, S., Linz, N., König, A., Langel, K., Pullen, H., Luz, S., Harrison, J. and Ritchie, C.W., 2022. Remote data collection speech analysis and prediction of the identification of Alzheimer's disease biomarkers in people at risk for Alzheimer's disease dementia: the Speech on the Phone Assessment (SPeAk) prospective observational study protocol. *BMJ open*, 12(3), p.e052250.

- Hacker, J., vom Brocke, J., Handali, J., Otto, M. and Schneider, J., 2020. Virtually in this together—how web-conferencing systems enabled a new virtual togetherness during the COVID-19 crisis. *European Journal of Information Systems*, 29(5), pp.563-584.
- Happyscribe., 2020. Privacy Policy | Happy Scribe. [online] Happy Scribe. Available at: <<https://www.happyscribe.com/privacy>> [Accessed 17 March 2022].
- Howlett, M., 2021. Looking at the 'field' through a Zoom lens: Methodological reflections on conducting online research during a global pandemic. *Qualitative Research*, p.1468794120985691.
- Identity Theft Resource Center. (2021). ITRC's 2021 Annual Data Breach Report. ITRC. Available at: <https://notified.idtheftcenter.org/s/2021-data-breach-report> (Accessed: 18 March 2022)
- Kvale, S. and Brinkmann, S., 2009. *Interviews: Learning the craft of qualitative research interviewing*. SAGE Publication Inc. California.
- Malterud, K., 2001. The art and science of clinical knowledge: evidence beyond measures and numbers. *The Lancet*, 358(9279), pp.397-400.
- McKenna, B., Myers, M.D. and Newman, M., 2017. Social media in qualitative research: Challenges and recommendations. *Information and Organization*, 27(2), pp.87-99.
- Merrett, F., 2006. Reflections on the Hawthorne effect. *Educational Psychology*, 26(1), pp.143-146.
- Microsoft. (2022a). Set default video permissions - Microsoft Stream. [online] Available at: <<https://docs.microsoft.com/en-us/stream/default-video-permissions>> [Accessed 20 May 2022].
- Microsoft. (2022b). Hardware requirements for Microsoft Teams - Microsoft Teams. [online] Docs.microsoft.com. Available at: <<https://docs.microsoft.com/en-us/microsoftteams/hardware-requirements-for-the-teams-app>> [Accessed 24 June 2022].
- Munsch, A., 2021. Millennial and generation Z digital marketing communication and advertising effectiveness: A qualitative exploration. *Journal of Global Scholars of Marketing Science*, 31(1), pp.10-29.
- National Cyber Security Centre. (2021). Alert: Further ransomware attacks on the UK education sector by cyber criminals. Online: GCHQ. Available at: <https://www.ncsc.gov.uk/files/NCSC-Alert-further-ransomware-attacks-UK-education-sector-20210604.pdf>
- Nicklin, L.L., Wilsdon, L., Chadwick, D., Rhoden, L., Ormerod, D., Allen, D., Witton, G. and Lloyd, J., 2022. Accelerated HE digitalisation: Exploring staff and student experiences of the COVID-19 rapid online-learning transfer. *Education and Information Technologies*, pp.1-26.
- Poorjam, A.H., Little, M.A., Jensen, J.R. and Christensen, M.G., 2019. Quality control in remote speech data collection. *IEEE Journal of Selected Topics in Signal Processing*, 13(2), pp.236-243.
- Riches, S., Elghany, S., Garety, P., Rus-Calafell, M. and Valmaggia, L., 2019. Factors affecting sense of presence in a virtual reality social environment: A qualitative study. *Cyberpsychology, Behavior, and Social Networking*, 22(4), pp.288-292.
- Taylor, P., Scholarios, D. and Howcroft, D., 2021. Covid-19 and working from home survey: preliminary findings.
- Tracy, S.J., 2010. Qualitative quality: Eight "big-tent" criteria for excellent qualitative research. *Qualitative inquiry*, 16(10), pp.837-851.
- University of Wolverhampton. (2022) FEHW Ethics Resources - University of Wolverhampton. Available at: <https://www.wlv.ac.uk/research/research-policies-procedures--guidelines/ethics-guidance/fehwh-ethics-resources/> (Accessed: 9 February 2022)
- VB-Audio. 2021. VoiceMeeter. (1.0.7.8). [Software]. [Accessed 18 March 2022].
- VB-Audio. (2022). VB-Audio VoiceMeeter. Available at: <https://vb-audio.com/Voicemeeter/>
- Zhang, C., Jepson, K., Lohfink, G. and Arvaniti, A., 2021. Comparing acoustic analyses of speech data collected remotely. *The Journal of the Acoustical Society of America*, 149(6), pp.3910-3916.
- Zoom. (2022b). Recording meetings locally and to the cloud. Available at: <https://support.zoom.us/hc/en-us/articles/360040771752-Recording-meetings-locally-and-to-the-cloud>. (Accessed 11 March 2022)
- Zoom. (2022a). Security Bulletin. Available at: <https://explore.zoom.us/en/trust/security/security-bulletin/>. (Accessed
- Zoom. (2022c). Zoom system requirements: Windows, macOS, Linux. Available at: <https://support.zoom.us/hc/en-us/articles/360040771752-Recording-meetings-locally-and-to-the-cloud>. (Accessed 11 March 2022)