Journal of the Economic Science Association (2024) 10:609–619
https://doi.org/10.1007/s40881-024-00165-7

EXPERIMENTAL TOOLS

MTVE: Magdeburg tool for video experiments

Dmitri Bershadskyy[1] · **Sunil Ghadwal**[1] · **Jannik Greif**[1]

Received: 24 February 2023 / Revised: 19 January 2024 / Accepted: 6 February 2024 /
Published online: 2 April 2024

Abstract

MTVE is an open-source software tool (citeware) that can be applied in laboratory and online experiments to implement video communication. The tool enables researchers to gather video data from these experiments in a way that these videos can be later used for automatic analysis through machine learning techniques. The browser-based tool comes with an easy user interface and can be easily integrated into z-Tree, oTree (and other experimental or survey tools). It provides the experimenters control over several communication parameters (e.g., number of participants, resolution), produces high-quality video data, and circumvents the Cocktail Party Problem (i.e., the problem of separating speakers solely based on audio input) by producing separate files. Using some of the recommended Voice-to-Text AI, the experimenters can transcribe individual files. MTVE can merge these individual transcriptions into one conversation.

Keywords Communication · Tool · Software

JEL Classification C80 · C88 · C90

1 Introduction

This paper introduces the Magdeburg Tool for Video Experiments (MTVE). MTVE is free citeware intended to assist researchers who want to capture audio and video data in laboratory or online experiments. By capturing the videos locally, MTVE avoids frequent issues, e.g., changing resolution rates, and enhances the high quality

✉ Dmitri Bershadskyy
dmitri.bershadskyy@ovgu.de

Sunil Ghadwal
sunil.ghadwal@ovgu.de

Jannik Greif
jannik.greif@ovgu.de

[1] Faculty of Economics and Management, Otto-von-Guericke University, Magdeburg, Germany

of transcription of communication. The tool is browser-based and compatible with typical software, e.g., z-Tree (Fischbacher, 2007) or oTree (Chen et al., 2016). All files (video, audio, transcriptions) are stored on the servers of the respective laboratory. Thus, MTVE enables better protection of the data of the participants.

Although communication has been investigated in experimental economics from early on (Isaac & Walker, 1988, 1991) even recent high-quality reviews on the implementation of communication in the laboratory do not yield unambiguous findings (Brandts et al., 2019). Instead, two insights can be made. First, if researchers are interested in one specific aspect of communication, it is valuable to restrict all parameters but the one of interest. This may refer to analyzing the effects of communication medium (Bochet et al., 2006; Brosig et al., 2003; Cason & Khan, 1999; Greiner et al., 2012, 2014; Isaac & Walker, 1988), direction (Cooper & Kühn, 2014), order (Cooper et al., 1989, 1992; Ottaviani & Sørensen, 2001), duration (Karagözoğlu & Kocher, 2016).

Second, unless there are precise reasons to restrict communication, free-form communication should be preferred as it is more externally valid. Yet, implementing free-form face-to-face communication in the lab leads to the question of how to analyze it (Brandts et al., 2019). A simple alternative is to use video communication. Conducting video communication enables researchers to record and later analyze the entire communication. This goes beyond the typical content analysis (Penczynski, 2019; Xiao & Houser, 2005) but may include analyzing facial expressions or voices as it is known that these are relevant for economic decisions (Antonakis et al., 2021; Bershadskyy, 2023; Centorrino et al., 2015; Hopfensitz & Mantilla, 2019).

To analyze videos from experiments automatically, researchers can either use proprietary software such as FaceReader (Serra-Garcia & Gneezy, 2021) or specially trained algorithms (Othman et al., 2019). Both approaches have different pros and cons.[1] However, both require high-quality data. This is where MTVE will help. In contrast to highly specialized software designed to record videos for research purposes, e.g., Noldus Viso[2] or Mangold VideoSyncPro,[3] MTVE is open-source and can be used without specialized hardware. While other companies are offering to design similar products, such solutions are costly.[4] An alternative would be to use video conference software, e.g., Zoom, Skype, or Big Blue Button (open source) as is done in some publications (Dulleck et al., 2017; Kachelmeier & Rimkus, 2022; Li et al., 2021). Still, such software is designed for conferencing and is not specialized in the requirements for analyzing communication in experiments, which we discuss in the following.

[1] The first approach applies always the same software and is simpler. Yet, the software is mostly a black box and was not trained on the specific data set. The second approach is likely to yield higher prediction rates yet requires more programming skills. Further, unless the code is published, it makes harder to reproduce the results.

[2] https://www.noldus.com/viso.

[3] https://www.mangold-international.com/en/products/software/video-recording-with-videosyncpro.html.

[4] According to our experience, prices demanded start at 45,000$.

The remaining article is structured as follows. In Sect. 2, we will describe the requirements for data from video conferences such that it can be analyzed through automatic tools. In Sect. 3, we explain how MTVE tackles these issues. Sections 4 and 5 focus on technical requirements and limitations respectively. Section 6 concludes.

2 Quality of video data analysis

In this section, we briefly discuss what is required to obtain an integrated analysis of video communication in experiments. We distinguish three major research factors: (i) video, (ii) audio, (iii) content and extend by a fourth factor—(iv) user-friendliness of the tool as it contributes to the replicability of experiments.

The most evident parameter of video data is resolution. High resolution is required to obtain an appropriate level of detail, e.g., participants' facial expressions. Yet, what is appropriate is mostly unclear. A good rule of thumb is to say the higher the resolution the better, given that it is always possible to decrease the resolution afterward as in Dudzik et al. (2021). What is equally important, is the framerate measured in frames per second (FPS). Recording at 60 FPS provides twice the data as at 30 FPS and increases the chances of algorithms detecting very short-lived movements. Similar to resolution, it is possible to decrease the FPS later.[5] Yet, for the analysis, it is important to keep FPS and resolution as constant as possible which normal videoconference tools (e.g., Zoom) cannot achieve as both parameters depend on the bandwidth.

Concerning audio, different parameters can be measured (e.g., volume, prosody). To measure these well, it is important to establish a silent surrounding to avoid acoustic disturbances, which is usually easy to achieve. The problem arises when more than one person takes part in the communication. This concerns the so-called "Cocktail Party Problem" (Cherry, 1953) which refers to the remarkable ability of humans to identify individual sources of acoustic input (e.g., voices) in a noisy environment. At the same time, it poses problems for neural networks (Haykin & Chen, 2005) making it difficult to separate speakers solely based on audio input (Ephrat et al., 2018). Therefore, it would be useful to solve this problem beforehand.

Referring to content, experimental economics has established a few different approaches summarized in Brandts et al. (2019). Independent of the chosen approach, the first step is to obtain the conversation as a text. This requires the experimenters to transcribe the text manually, outsource it to transcription companies, or use speech-to-text software. Yet, such software also suffers from the Cocktail Party Problem.

Finally, we consider user-friendliness from the perspective of replicability of experiments. The easier it is to replicate an experiment, the more likely is a

[5] Please note, that we do not consider other parameters that influence the quality of recordings (e.g., lighting and the general surrounding of participants in the laboratory). Instead, we focus on the technical aspects.

replication. Further, as video conferences handle personalized data (e.g., faces), it is preferable not to share participants' data with other companies (e.g., Skype, Zoom). Therefore, the goal is an open-source tool with an intuitive interface that is operable for normal laboratory and online experiments and where the data does not leave the digital space of the laboratory.

3 The structure of MTVE

MTVE consists of three apps (Meetings App, Video App, and Transcription App). The Meetings App is used to organize the rooms. When participants join the created room, the Video App starts. After communication ends, the experimenters can start transcribing data using the Transcription App. Altogether they tackle the issues discussed in Sect. 2. In this section, we briefly discuss the structure of MTVE and refer to a more detailed description on GitHub (https://github.com/MaXLab-OVGU/MTVE).

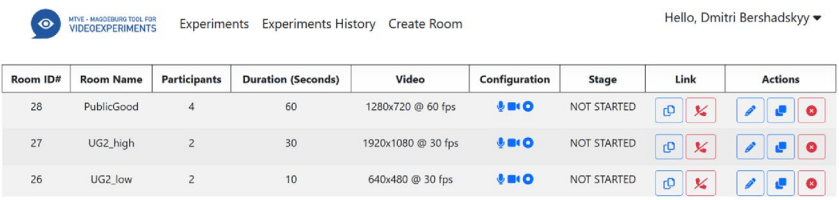
Using the simple user interface (UI) of MTVE (see Fig. 1a), experimenters can create a room with a few clicks and configure different features (see Fig. 1b). After the communication room is created, the experimenter receives a simple link that can now be implemented into the experiment. In oTree this will start a new browser tab and in z-Tree it will start a new browser window. Since the rooms can be reused, nothing has to be changed between the sessions. Further, we integrated the option for the experimenters to close the room remotely.

All that is required to do, is that laboratory managers install MTVE on their servers. MTVE comes along with online documentation[6] that explains how to do so. Laboratory managers become administrators of MTVE on their local server and guide the experimenters on how to register an account. MTVE provides a simple registration procedure where the experimenters create their accounts using e-mail and password. After confirmation, the experimenter can log in and create their experiments. The laboratory manager as the administrator is permitted access to all accounts, enabling them to support novice experimenters.

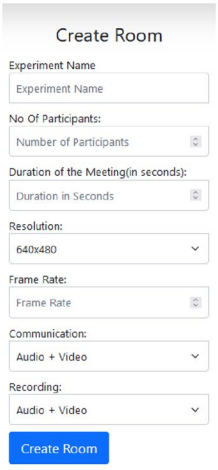
From the perspective of experimental participants, all they see is a waiting screen (see Fig. 1c) as long as the predetermined group size is not reached. The communication starts (see Fig. 1d) once every participant has joined the room. The tab can be closed automatically using z-Tree or oTree without the need for participants' engagement.

The most important feature of MTVE is how it avoids the Cocktail Party Problem. Before the audio–video signal is uploaded to the server, combining all audio–video streams into a joint communication, MTVE captures an additional copy of the signal that is saved as a separate file and is automatically transferred to the laboratory server separately. This approach (see Fig. 2) does not only solve the Cocktail Party Problem but further stabilizes the quality (e.g., resolution) of the individual video files.

[6] https://github.com/MaXLab-OVGU/MTVE.

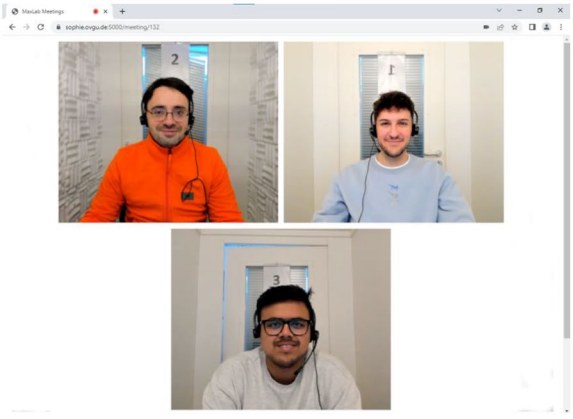


(a) Overview of the user interface, including a summary of existing rooms, current status, and the option to close the room



(b) Creating room as experimenter

(c) Waiting screen for participants until everyone joined



d) Communication in a browser tab for three individuals

Fig. 1 User interface

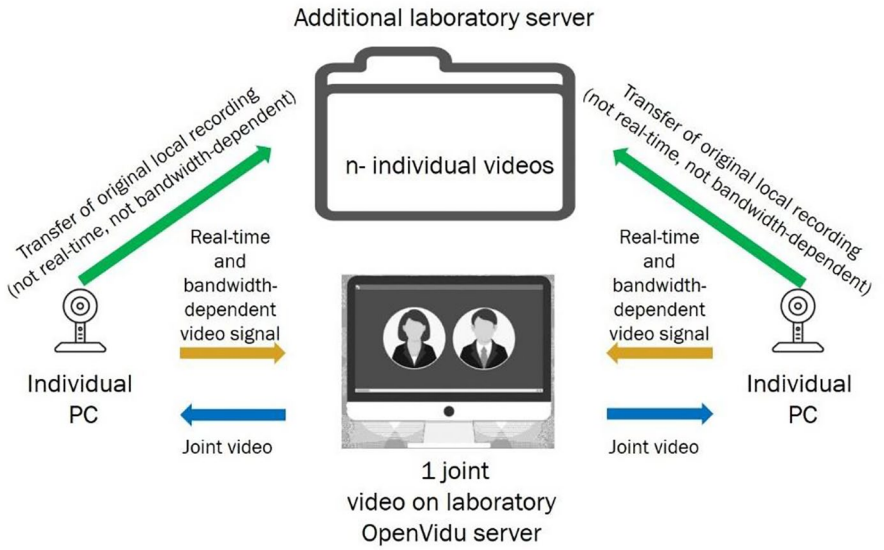


Fig. 2 MTVE's approach to the cocktail party problem. In this example, two participants are communicating with each other. To provide normal communication, the signal goes to the OpenVidu server which sends the joint video back to both participants. Independent of this process, MTVE captures a local recording of the individual videos in high quality and sends them to a second server. This yields $n+1$ videos for n-person communication

These individual files can now be transcribed using a variety of different voice-to-text AI models (e.g., Whisper, VOSK[7]). The quality of transcription depends on the audio quality (e.g., headphones, background noises) and the model itself. We provide one possible solution within our tool, yet stress that researchers could try different models and foremost should recheck the transcriptions.[8] Still, even if transcribing individual voices worked well, it is necessary to merge individual transcriptions. Here, MTVE offers a solution that merges transcriptions of individual video files into one CSV, generating the original chat structure of communication.

Altogether, MTVE is easy to use for all groups of users (laboratory managers, experimenters, and experimental participants) and yields the experimenters a standardized set of communication data that can be analyzed. This leads to the remaining question of whether the technological benchmark for the laboratories is sufficiently low.

[7] VOSK: https://alphacephei.com/vosk/models; Whisper: https://github.com/openai/whisper.

[8] This implies that there still has to be a human involved. Yet, we argue that simply checking whether the transcription is accurate requires much less time than actually transcribing the discussion.

4 Technical requirements and license

In our discussion of the technical requirements of MTVE, we focus on three parts: Meetings App (UI), Video App, and Transcription Utility. We stress that the requirements for the first ones are moderate yet can be high for the last one. The next paragraphs shall discuss these issues in more detail and we provide additional information on GitHub.

The UI consists of an app created in Python with the Django framework. Both can be considered very common components and do not impose major challenges on a typical experimental laboratory. The same is true for the MySQL server used to store and retrieve the data. Second, the video app consists of an express.js server created and ran using Node js, and an OpenVidu instance running on a separate server to handle video streams. The Meetings App and the Video App connect to the same MySQL server to store and retrieve data. Concerning technical equipment, we indicate that, when researchers run too many rooms simultaneously, this may affect the quality of the recordings (e.g., saved videos may be shorter than set). This issue depends on the server capacity. We provide performance tests for certain reasonable combinations of room sizes and number of rooms on GitHub to give researchers an estimate on the required server setup. Still, we advise researchers to run a few tests on their own.

The largest technical requirement comes from the transcription tool. The first step consists of a Python module Pydub (FFMPEG) to convert video files to wav files. In the second step, it applies a predefined transcription model. These models, however, can impose more or less requirements. For instance, Whisper requires up to 10 GB VRAM while VOSK models require up to 4.4 GB for the German language model. We stress, that researchers can implement the model of their choice. In the end, the script that puts the separate texts together depends on the filenames that are generated implying that these shall not be changed.

MTVE is open-source, licensed under an adaptation of the MIT license.[9] We ask researchers to cite this paper when using MTVE for academic or other publications. The source code for MTVE can be downloaded for free from GitHub. Contributions and improvements to the source code are welcome and should be submitted using GitHub, too.

5 Limitations

As of the current version (MTVE 1.1), the tool has certain limitations, which we present in this chapter. It is our goal to improve on these. We refer to the GitHub project[10] for possible updates and pull requests.

[9] https://github.com/MaXLab-OVGU/MTVE/blob/main/LICENSE.
[10] https://github.com/MaXLab-OVGU/MTVE.

While MTVE supports resolution up to 3840×2160 and 60FPS it is currently not possible to combine both. Recording with 60 FPS is only possible for HD resolution and lower.

Further, we stress again that the quality of the transcription does not depend on MTVE but on the actual model chosen. We highlight that whatever model is chosen, has to produce time stamps, so MTVE can merge the individual text files into one.

MTVE was tested on Chrome and Firefox. To guarantee a simultaneous entry to the conversation for all participants, it is essential to configure the browser in such a way that the application of the webcam does not need further acceptance from the user. Otherwise, some participants join the conversation but cannot be seen by others until they accept manually.

6 Conclusion

All in all, MTVE enables experimenters to gather communication data from experiments with video conferences in an integrated approach. The tool is easily implementable in z-Tree, oTree, and other software. It saves high-quality videos. Further, applying a simple trick, it avoids the Cocktail Party Problem and enables more sophisticated analysis of the voices. MTVE comes with a simple user interface for the experimenter and comparatively low technical requirements for the laboratory. The tool is open-source (citeware) and can easily be adapted to local servers of any laboratory. Doing so enables researchers to keep subjects' data on private servers and strengthens local data protection regulations.

Appendix A

Use Case 1

MTVE is used in z-Tree to enable communication between the participant and experimenter to record participants' answers in the seminal dice rolling experiment (Fischbacher & Föllmi-Heusi, 2013). Here, after rolling the dice, the participants start the communication and inform the experimenter of what they claim to have rolled.

Use Case 2

MTVE is used in oTree to analyze the experiment from Belot and van de Ven (2017) where there are two types of players. Player 1 has the incentive to lie about the true color of their card (red or black) and Player 2 has the incentive to detect the lie. Per session, there are five Players 1 and five Players 2. Using the round robin scheme all Players 1 and Players 2 are matched. Player 1 indicates the color to Player 2 using MTVE in a predefined time of 10 s. After that, the communication stops.

Use Case 3

MTVE is used in z-Tree to allow repeated communication between two subjects. Further, it is used to allow the experimenter to close the communication room remotely when needed.

Acknowledgements We acknowledge financial support from the German Research Foundation (DFG) through the project number 468478819. The authors would like to thank the editor Lionel Page and two anonymous reviewers for their comments and suggestions. Moreover, we acknowledge valuable advices on functionalities of MTVE from: Ayoub Al-Hamadi, René Degenkolbe, Laslo Dinges, Marc-André Fiedler, Nina Ostermaier, Myra Spiliopoulou, and Joachim Weimann. We further acknowledge the support of Adrija Ghosh, Raviteja Sutrave, and Sourima Dey during a scientific student project.

Funding Open Access funding enabled and organized by Projekt DEAL.

References

Antonakis, J., d'Adda, G., Weber, R. A., & Zehnder, C. (2021). Just words? Just speeches? On the economic value of charismatic leadership. *Management Science*. https://doi.org/10.1287/Mnsc.2021.4219

Belot, M., & van de Ven, J. (2017). How private is private information? The ability to spot deception in an economic game. *Experimental Economics, 20*(1), 19–43. https://doi.org/10.1007/s10683-015-9474-8

Bershadskyy, D. (2023). Reverberation effect of communication in a public goods game. *PLoS ONE, 18*(2), e0281633. https://doi.org/10.1371/JOURNAL.PONE.0281633

Bochet, O., Page, T., & Putterman, L. (2006). Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior & Organization, 60*(1), 11–26. https://doi.org/10.1016/J.JEBO.2003.06.006

Brandts, J., Cooper, D. J., & Rott, C. (2019). Communication in laboratory experiments. In A. Schram & A. Ule (Eds.), *Handbook of research methods and applications in experimental economics* (pp. 401–426). Edward Elgar Publishing.

Brosig, J., Weimann, J., & Ockenfels, A. (2003). The effect of communication media on cooperation. *German Economic Review, 4*(2), 217–241. https://doi.org/10.1111/1468-0475.00080

Cason, T. N., & Khan, F. U. (1999). A laboratory study of voluntary public goods provision with imperfect monitoring and communication. *Journal of Development Economics, 58*(2), 533–552. https://doi.org/10.1016/S0304-3878(98)00124-2

Centorrino, S., Djemai, E., Hopfensitz, A., Milinski, M., & Seabright, P. (2015). Honest signaling in trust interactions: Smiles rated as genuine induce trust and signal higher earning opportunities. *Evolution and Human Behavior, 36*(1), 8–16. https://doi.org/10.1016/J.EVOLHUMBEHAV.2014.08.001

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance, 9*, 88–97. https://doi.org/10.1016/j.jbef.2015.12.001

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America, 25*(5), 975–979. https://doi.org/10.1121/1.1907229

Cooper, D. J., & Kühn, K. U. (2014). Communication, renegotiation, and the scope for collusion. *American Economic Journal: Microeconomics, 6*(2), 247–278. https://doi.org/10.1257/MIC.6.2.247

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1989). Communication in the battle of the sexes game: Some experimental results. *The RAND Journal of Economics, 20*(4), 568. https://doi.org/10.2307/2555734

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1992). Communication in coordination games. *The Quarterly Journal of Economics, 107*(2), 739–771. https://doi.org/10.2307/2118488

Dudzik, B., Columbus, S., Hrkalovic, T. M., Balliet, D., & Hung, H. (2021). Recognizing perceived interdependence in face-to-face negotiations through multimodal analysis of nonverbal behavior. In *ICMI 2021—Proceedings of the 2021 international conference on multimodal interaction* (pp. 121–130). https://doi.org/10.1145/3462244.3479935.

Dulleck, U., He, Y., Kidd, M. P., & Silva-Goncalves, J. (2017). The impact of affirmative action: Evidence from a cross-country laboratory experiment. *Economics Letters, 155*, 67–71. https://doi.org/10.1016/j.econlet.2017.03.010

Ephrat, A., Mosseri, I., Lang, O., Dekel, T., Wilson, K., Hassidim, A., Freeman, W. T., & Rubinstein, M. (2018). Looking to listen at the cocktail party: A speaker-independent audio-visual model for speech separation. *ACM Transactions on Graphics, 37*(4), 11. https://doi.org/10.1145/3197517.3201357

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10*(2), 171–178. https://doi.org/10.1007/s10683-006-9159-4

Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise—An experimental study on cheating. *Journal of the European Economic Association, 11*(3), 525–547. https://doi.org/10.1111/JEEA.12014

Greiner, B., Caravella, M., & Roth, A. E. (2014). Is avatar-to-avatar communication as effective as face-to-face communication? An ultimatum game experiment in first and second life. *Journal of Economic Behavior & Organization, 108*, 374–382. https://doi.org/10.1016/J.JEBO.2014.01.011

Greiner, B., Güth, W., & Zultan, R. (2012). Social communication and discrimination: A video experiment. *Experimental Economics, 15*(3), 398–417. https://doi.org/10.1007/s10683-011-9305-5

Haykin, S., & Chen, Z. (2005). The cocktail party problem. *Neural Computation, 17*(9), 1875–1902. https://doi.org/10.1162/0899766054322964

Hopfensitz, A., & Mantilla, C. (2019). Emotional expressions by sports teams: An analysis of World Cup soccer player portraits. *Journal of Economic Psychology, 75*, 102071. https://doi.org/10.1016/j.joep.2018.04.008

Isaac, R. M., & Walker, J. M. (1988). Communication and free-riding behavior: The voluntary contribution mechanism. *Economic Inquiry, 26*(4), 585–608. https://doi.org/10.1111/j.1465-7295.1988.tb01519.x

Isaac, R. M., & Walker, J. M. (1991). Costly communication: An experiment in a nested public goods problem. In T. R. Palfrey (Ed.), *Laboratory research in political economy* (pp. 269–286). University of Michigan Press.

Kachelmeier, S. J., & Rimkus, D. (2022). Does seeking audit evidence impede the willingness to impose audit adjustments? *The Accounting Review, 97*(7), 269–293. https://doi.org/10.2308/TAR-2019-0613

Karagözoğlu, E., & Kocher, M. G. (2016). *Bargaining under time pressure* (No. 5685; CESifo working paper series). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2727156.

Li, J., Leider, S., Beil, D., & Duenyas, I. (2021). Running online experiments using web-conferencing software. *Journal of the Economic Science Association, 7*(2), 167–183. https://doi.org/10.1007/S40881-021-00112-W

Othman, E., Saxen, F., Bershadskyy, D., Werner, P., Al-Hamadi, A., & Weimann, J. (2019). Predicting group contribution behaviour in a public goods game from face-to-face communication. *Sensors, 19*(12), 2786. https://doi.org/10.3390/s19122786

Ottaviani, M., & Sørensen, P. (2001). Information aggregation in debate: Who should speak first? *Journal of Public Economics, 81*(3), 393–421. https://doi.org/10.1016/S0047-2727(00)00119-5

Penczynski, S. P. (2019). Using machine learning for communication classification. *Experimental Economics, 22*, 1002–1029. https://doi.org/10.1007/s10683-018-09600-z

Serra-Garcia, M., & Gneezy, U. (2021). Mistakes, overconfidence, and the effect of sharing on detecting lies. *American Economic Review, 111*(10), 3160–3183. https://doi.org/10.1257/AER.20191295

Xiao, E., & Houser, D. (2005). Emotion expression in human punishment behavior. *Proceedings of the National Academy of Sciences of the United States of America, 102*(20), 7398–7401. https://doi.org/10.1073/pnas.0502399102