

AN OPEN SOURCE VIRTUAL REALITY TRAINING FRAMEWORK FOR THE CRIMINAL JUSTICE SYSTEM

HERMANN BARBE 

JÜRGEN L. MÜLLER

BRUNO SIEGEL

PETER FROMBERGER

Clinic for Psychiatry and Psychotherapy, Forensic Psychiatry, Human Medical Center Göttingen

We developed an open-source training framework to practice conversation skills in a controlled and immersive virtual reality (VR) environment. Virtual characters with different biographies were developed with which a conversation using natural language is possible. The virtual characters integrate a dialog management system (ChatScript) to provide different biographical memories. Natural language processing for the German language is integrated by using Kaldi, an open-source speech recognition toolkit. As the framework allows for interchangeable content there are many different possible application cases to apply within the criminal justice system. The VR frameworks code is available under an open-source license. In this article, an overview of the framework's functionality is given as well as an outlook on possible areas of application. Statements about user acceptance and usability cannot yet be made, as relevant data have first to be gathered through a concrete application case.

Keywords: criminal justice system; virtual reality training; forensic psychiatry; chatbot; virtual standardized patient; forensic conversation training; open source

The main advantage of virtual reality (VR) technology is the ability to create immersive virtual environments that feel authentic and plausible to the user. The concept of immersion describes the amount and quality of sensory stimulation a virtual environment can create (Bohil et al., 2011). It is thus not a psychological state, but a description of a VR system's properties. VR systems can create immersion by presenting specific stimuli, which are difficult to discriminate from real, nonvirtual stimuli (Ryan et al., 2019). This can, for example, be done by a combination of realistic visual environments, acoustic settings, and

AUTHORS' NOTE: *The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. This study was funded by the German Federal Ministry of Education and Research (BMBF, Funding number: 01SR1703C). The funder has not, and will not, participate in the study design, data collection, management, analysis, and interpretation of data, or writing of reports. Correspondence concerning this article should be addressed to Hermann Barbe, Clinic for Psychiatry and Psychotherapy, Forensic Psychiatry, Human Medical Center Göttingen, Rosdorfer Weg 70, Göttingen, 37081, Germany; e-mail: hermann.barbe@med.uni-goettingen.de.*

CRIMINAL JUSTICE AND BEHAVIOR, 2023, Vol. 50, No. 2, February 2023, 294–303.

DOI: 10.1177/00938548221124128

Article reuse guidelines: sagepub.com/journals-permissions



© 2022 International Association for Correctional and Forensic Psychology

head tracking, where the scenery is moving along with the rotation of the user's head. The degree of immersion can be manipulated by either adding (or removing) visual or acoustic stimuli or by increasing (or decreasing) the fidelity of the presented stimuli, by using a higher resolution head-mounted display (HMD). The ability to create immersive virtual environments is a valuable addition to common psychological and psychiatric research methods, assessment tools, and treatment possibilities in the forensic psychiatric field (Barbe et al., 2020; Sygel & Wallinius, 2021).

But VR applications and environments are not limited to a mere visual presentation of certain stimuli and situations, which participants can only experience passively as it has been done successfully in virtual reality exposure therapy (VRET; for an overview see Emmelkamp & Meyerbröker, 2021). They can also be interactive and used as training tools. A recent meta-analysis compared different forms of extended reality (XR)-based trainings—which VR is a part of—with traditional trainings (Kaplan et al., 2020). As XR-based training methods were found to be as effective as traditional techniques considering the training outcome, the authors suggest that XR-based training applications have promising potential, especially in areas, where traditional trainings are difficult to establish.

One area, which could utilize these advantages and the potential is the field of forensic psychiatry, which comes with the inherent demand for simulating high-risk social scenarios, for example, assessing the risk of violent offenders, as well as situations, which would raise ethical issues in real life, such as exposing sexual offenders to potential victims (for an overview on current VR tools utilizing this feature see Sygel & Wallinius, 2021). In conjunction with the high ecological validity due to the similarity to real-life situations, VR applications could therefore serve as a standardized training and therapy tool for forensic inpatients and offenders, which can be applied without endangering others (Fromberger et al., 2018). Fromberger et al. (2014) also implicated a possible implementation of VR-training tools for the education of forensic staff. One exemplary suggestion by Fromberger was the practice of professional handling and treatment of aggressive (forensic) inpatients, as VR provides the opportunity to train in a safe virtual environment.

Another important domain where interactive VR training applications can be utilized in the criminal justice system is the training of interpersonal communication skills as communication or interviewing skills are essential skillset in many different professions. Be it in examining witnesses in police investigations, taking the medical history of a patient, working as a psychotherapist with (mentally ill) offenders, or creating a forensic report. To the best of our knowledge, VR training tools to practice conversations in the context of the criminal justice system have not been established yet. Nevertheless, there have been some experiences with non-VR interview trainings in medical and psychiatric contexts.

Practicing interviewing skills in medical or psychological clinical education is commonly conducted with so-called human standardized patients (HSP), which have been used since the 1960s (Talbot et al., 2012). HSPs are trained actors who can portray different types of individual respondents (patients, witnesses, etc.) consistently throughout multiple interviews to simulate real-life interview situations. Studies showed that those trainings become particularly effective when students were given immediate feedback by a supervisor (Lamb et al., 2010). One important limitation of HSPs is that they are mostly limited to healthy appearing adults in their portrayal so interviewing children, elderly people, or person with impairments is hardly possible (Talbot & Rizzo, 2019).

Contemporary training designs aim to recreate such interview situations realistically with so-called virtual standardized patients (VSP), allowing trainees to practice their questioning technique and communication skills repeatedly with the help of virtual characters and avatars.¹ The benefit of such approaches is flexibility and cost-effectiveness (Wang et al., 2016), as implementation and training become independent of professional actors. Studies that compared human-avatar with human–human conversation showed that the general conversation techniques do not differ significantly implying a transferability of possible training gains to real-life situations (Heyselaar et al., 2017). An exemplary implementation of Maicher et al. (2017) combined a two-dimensional representation of a VSP on a computer screen with a chatbot-based dialog management system to help medical students to practice initial medical interviews. Participants could ask their question either using natural language or typed-in text. Their task was to gather a correct medical history using appropriate questions. In their study, 76.9 % of the 141 participating students correctly diagnosed the presented case. The authors concluded that the inclusion of a chatbot-based dialog management system was highly suitable to create VSP with realistic and adequate response behavior.

Building upon the VSP research as a training tool for medical staff, we introduce an open-source training framework for the criminal justice system. The framework not only utilizes chatbot technology to create realistic conversation experiences, but it also adds the ability to conduct conversations in a virtual immersive environment. In comparison to two-dimensional applications (e.g., Maicher et al., 2017), there are several advantages to high immersive virtual conversation scenarios. They raise the ecological validity by immersing the trainee in the virtual situation while also giving investigators more control over distractions during the training (Parsons & Rizzo, 2019). Finally, it is much easier to change the environmental context of a conversation (e.g., interviewing room, prison, courtroom) by adjusting the virtual setup (Bordnick & Washburn, 2019). Most importantly for applications within the criminal justice system and the work with forensic inpatients or offenders, VR-based training frameworks allow to train skills in a safe environment without endangering others in the case of mistakes by the trainee (Fromberger et al., 2014).

In the following, a technical overview of the framework’s logic and functionality is given by focusing on the description of the used software and possible cases of application. In summary, the framework allows practicing multiple conversations with virtual characters in a virtual environment using natural language. Conversations are being automatically analyzed based on desirable or undesirable techniques so that trainees can get immediate and standardized feedback after each interview. The virtual characters and the environment are interchangeable, allowing for implementation in different contexts of the criminal justice system.

VR-TRAINING FRAMEWORK

The development of the VR-training framework consequently follows the open-source approach, avoiding the usage of proprietary software. Thus, all software components either already use an open-source license or are written by the authors and are published under an open source license. The framework’s code is available on GitLab (hosting platform for software development) under the GNU General Public License 3 (“GPL v3”) and can be downloaded using the following link: https://gitlab.gwdg.de/barbe/vr_conversation_training.

DESIGN

The VR-training framework consists of five main components: (a) a web server that functions as a data storage, serves the web frontend, is used as a data management tool, and provides web-based questionnaires and training feedback to trainees immediately after finishing a training trial; (b) an operator application used to organize the experiment to collect operator input and to manage the data transfer to the server; (3) a conversation engine (“CE”) that enables conversation using natural language between trainee and virtual characters; (4) a virtual environment, where trainees interact with virtual characters; and (5) a speech-to-text framework, to allow freely spoken questions.

WEB SERVER

To simplify data management, a Linux-based server (Ubuntu, <https://releases.ubuntu.com>) is used as the central server facilitating data management, running a web framework, and allowing for the parallel usage of multiple training setups. The server consists of three main parts: a web framework (Django; <https://www.djangoproject.com/>), a web API (Django REST Framework, DRF; <https://www.django-rest-framework.org/>) and a database (SQLite; <https://www.sqlite.org>). The high-level Python web framework Django is the basis for the operator frontend and provides the backend for data management. The DRF allows the usage of the database via web APIs.

The training management frontend’s user interface (UI) was designed with Bootstrap (<https://getbootstrap.com/docs/4.0/getting-started/introduction/>). It is used to carry out different tasks related to the training simulation: (a) creation of a new trainee, (b) vignette² and questionnaire presentation, (c) visualization of individualized feedback, and (d) data export in a ready-to-use format for statistical programs. The main advantages of the web-based training management frontend are the capability to provide web-based questionnaires and automatized individual training feedback that is accessible via browser by any computer connected to the same local network.

The integrated feedback system automatically generates individual feedback for each training trial. It is written in Python allowing us to easily adapt the algorithm for different cases of application. Currently, the integrated feedback system can show the proportion as well as examples of predefined good and bad snippets of transcribed speech that occurred during the last virtual conversation. Furthermore, it can provide hints to avoid pitfalls in the next conversation. The feedback is balanced over multiple conversations by a progressive analysis of already given individual feedback to lower the number of repetitive snippets or hints. Furthermore, it can be specified at which timepoint a questionnaire has to be filled out or when additional information should be provided to the user. Currently, two time points (before or after the VR training) are implemented.

Django provides a built-in authentication and permission system based on different user roles and groups. This makes it possible to define different user roles with individual permissions to minimize the risk of accidental data manipulations. The following groups are defined: (a) administrators, who have the permission to export data, (b) training operators, who are allowed to create new trainees and training feedback, and (c) trainees, who are only allowed to enter questionnaires and view the training feedback. The structure of the relational database has been developed to ensure data integrity and to prevent the creation of

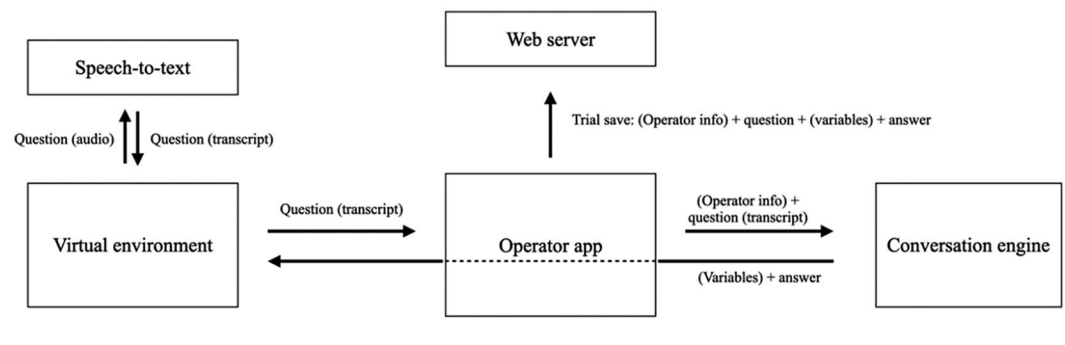


Figure 1: Information Flow of VR-Training Framework

Note. Operator application is responsible for the message transfer between the conversation engine and the virtual environment. Using the application, operators can add additional information to the question transcript and broadcast both (information + transcript) back to the conversation engine. VR = virtual reality.

duplicates or overwriting already created data. Django provides all the necessary tools to manage the relational database out of the box.

OPERATOR APPLICATION

The operator application was written in node.js (Node.js; <https://nodejs.org>) using JavaScript, ran using the open source framework Electron (OpenJS Foundation, Electron; <https://www.electronjs.org>), and communicates with the database via web API as described above. Each VR training is started using the operator application. Within this application, the operator configures the training setup (e.g., which user performs the training, which virtual character is used etc.). To achieve this, the application makes a request to the web server, showing the (remaining) possible options for each subject and thus ensuring data integrity. After the conversation is started, the application functions as a control center between the conversation engine, the virtual environment, and the database. See Figure 1 for a schematic depiction of information flow.

Accordingly, an exemplary trial looks as follows: after a trainee asks a question, it is first being transcribed by the speech-to-text module. The operator application then receives the transcribed question in text form. At this stage, it is possible to allow the operator to apply additional information for the conversation engine. This is realized by a pop-up window that opens when the operator application receives a new transcript. The operator can then either add details about the transcripts content or about the appropriate virtual character's response behavior, for example, aggressive or sad behavior. If there is no need for additional information, transcripts can also be sent directly to the conversation engine. The conversation engine analyzes the question (and operator information if provided), finds the appropriate answer, and sends it back as a text to the operator app. The question-answer-combination—as well as additional variables that are calculated inside the conversation engine—is then sent to the web server and displayed inside the operator application in the form of a written chat record, allowing the operator to keep track of the development of the conversation over the course of the experiment. The operator application also saves the question-answer-combination in the database and sends the answer to the virtual environment, which presents the answer to the trainee

in spoken form. Conversations can be stopped either manually by the operator or automatically according to a pre-set time limit. In both cases, the operator application sends an end message to the virtual environment to stop the simulation.

CONVERSATION ENGINE

The text-based dialog management system ChatScript (Wilcox & Wilcox, 2013) was used to analyze the transcribed question. It enables the training framework to produce adequate answers based on a predefined virtual character's biography given the conversation's context. ChatScript was already successfully applied in a similar role in training studies for medical students with virtual standardized patients (Maicher et al., 2017, 2019). The studies showed that the ChatScript-based patients produced a high rate of adequate answers to the students' questions and thus made it possible to simulate realistic response behavior. The rate of appropriate answers is an important factor in such interactive training designs as frequent inappropriate answers can quickly elicit a sense of unrealistic response behavior (Johnsen et al., 2005).

ChatScript has extensive pattern-matching capabilities to analyze given input which allows one to break down the intent of a question or statement in a few characteristic keyword patterns. In consequence, multiple different questions or statement variations, with the same intent, can be summarized allowing adequate responses to a wide range of different statements. ChatScript allows structuring the virtual characters' knowledge in so-called topics. These are collections of the aforementioned patterns that involve a certain common theme, like hobbies or friends. Topics are divided by characteristic keywords. That way responses keep track of context and thus, the conversation can switch to another topic, if a certain keyword is used or all answers inside a topic are already elicited. The conversation engine furthermore keeps track of the conversation course and determines adequate emotional reactions and response behavior of the virtual character considering the context. This is possible by the implementation of variables and functions, similar to other common programming languages. Variables are utilized to perform two different tasks. The first is to supply descriptive information about the current state of the conversation for each trial. The second task is to establish a degree of self-awareness for the virtual character, for example, count the number of times a specific type of question has been asked. If a certain condition is met (e.g., the subject surpassed a number of specific question types) the virtual character can react with appropriate predefined behavior. The necessary behavior is added to the descriptive information as a variable and thus available for the virtual environment.

VIRTUAL ENVIRONMENT

The virtual environment was set up in Unity (Unity Technologies, Unity3D, <https://unity.com>). Unity is a game engine supporting a range of different VR hardware. Its usage is free for nonprofit purposes. It provides a WYSIWYG ("What you see is what you get")-Editor for 3D environments, and more complex functions can be added using the C# programming language. The training framework comes with a number of different emotional expressions that can be applied to human character models to provide realistic communication behavior of the virtual characters. The emotional expressions were captured using motion-capturing techniques and could be distinguished by different body postures. The set consists of five

different emotional expressions: (a) neutral, (b) sad, (c) serious, (d) disappointed, and (e) relieved. The emotional expressions are organized using an animation controller. The controller receives the relevant emotional expression information from the conversation engine's recent answer. If there is a change of emotion, the controller cross-fades the animation using a flowing transition over the previous one. As long as no such information is received, the current animation will repeat in a loop.

The virtual characters spoken answers were implemented using Google Cloud text-to-speech (<https://cloud.google.com/text-to-speech>). For this purpose, the German WaveNet voice types were used and different sounding virtual characters were achieved by manipulating pitch and speed of speech. The sound files are managed inside the Unity project as well: Whenever Unity receives the information about the text-based answer from the conversation engine, Unity finds the corresponding audio file and plays it over the HMD's headphones. The advantage of using artificially generated spoken answers is that they bypass the ethical problem of recording real children talking about sensitive topics like sexual abuse or physical violence. In addition, it is a highly efficient way of generating the large body of spoken answers necessary to create the illusion of natural conversation.

Before the subject's question can be transcribed in the speech-to-text module, it has to be recorded first. This was solved by using the HMD's controller and microphone. Whenever the user pulls the controllers' trigger, the sound is being recorded over the microphone. After the subject releases the trigger, the audio recording is stopped, a WAV file is created, and sent to the speech-to-text module for transcription. To practice the timing of this solution, we implemented a tutorial to teach users to start speaking after the trigger pull and to stop speaking before the trigger release. Generating a correct recording is essential for a clean transcription and further analysis inside the conversation engine.

SPEECH-TO-TEXT

To transcribe the recorded audio file into text form to be analyzed by the conversation engine we implemented a speech-to-text module. Following open-source technical regulations, Kaldi (Povey et al., 2011), a speech recognition toolkit was utilized for speech-to-text processing in real-time. Kaldi runs on a virtual Ubuntu machine. Communication with Kaldi is established by setting up a server within the virtual machine, which the Unity-run virtual environment then automatically connects to on startup. When the server receives an audio file, Kaldi transcribes it and sends back the corresponding transcription text. As the underlying language model, we implemented a pretrained German model by the Language Technology Group, Universität Hamburg (UHH; Milde & Köhn, 2018). In their article, the authors analyzed the quality of transcription by calculating the word error rate³ (WER) and found a WER of around 14.4%. It contains a vocabulary of 400,000 words and was trained on 630 hr of audio data (model state: 05.03.2019).

DISCUSSION

The presented open-source VR-training framework allows for simulating a virtual scenario, where trainees or professionals working in the criminal justice system as well as forensic patients can practice their conversation skills in a safe environment. Conversations are conducted using free speech and natural language resulting in realistic conversational

situations. The virtual characters also contribute to a realistic experience as they can answer adequately, given the context as well as look and behave like a real person. Trainees get standardized reports considering their interview techniques after each conversation, supporting the steady improvement in their training results.

The combination of immersive virtual environments and chatbot technology for adequate responses could be the next step toward a new generation of training applications to practice conversations in the criminal justice system. This approach can improve the standardization of conversation trainings as well as the level of realism concerning the similarity between the training scenario and real-world situations. The main advantages of the framework are that it is not limited to one particular scientific context due to the inclusion of interchangeable content and that the code is published on an open-source code sharing platform. That means that interested research groups or practitioners can focus mainly on the development of the content-related aspects of a training conversation context—environment (e.g., interviewing or patient room, playground, prison), virtual character models, and biographies (e.g., adults, children, patients, witnesses)—to implement them in the existing framework. This way we hope to simplify the realization of such applications in the domain of criminal justice. With the publication of the framework, we also hope to speed up the development process for future applications, as everyone can download the code for free and start implementing interesting ideas and content.

The possibility of practicing conversation skills is a very important topic in the field of forensic psychiatry and psychology both for practitioners and for inpatients and clients. As a result, there are many different contexts where a VR-training framework could be applicable such as practicing initial interviews and counseling for therapists and counseling centers; interviewing witnesses or assessment and training for sexual offenders in virtual critical situations; and interrogation training for law enforcement investigators.

LIMITATIONS

One limitation at the moment is the lack of data to ensure important usability features of the VR training. In addition to simulator sickness, which can cause participants to experience symptoms such as dizziness or malaise, these include subjects' assessments of how realistic they felt the VR training was. Both aspects should be recorded and analyzed in future implementations alongside the variables under investigation.

Another limiting factor is that the published code version for the VR-training framework has still some experimental characteristics that require a basic understanding of programming languages to setup properly and implement custom content files. It also needs to be mentioned that the VR-framework in the published version does not (yet) have encrypted data transmission implemented. That means if an implementation in an open or public network is indented, this has to be considered by customizing the frameworks' communication interfaces.

AUTHOR CONTRIBUTIONS

H.B., P.F., B.S., and J.L.M.: conceptualization. P.F. and J.L.M.: funding acquisition. H.B., P.F., B.S., and J.L.M.: methodology. H.B.: visualization, writing—original draft. P.F. and J.L.M.: project administration. H.B., B.S., and P.F.: software. P.F. and B.S.: supervision. P.F., B.S. and J.L.M.: review and editing. All authors have read and approved the manuscript.

ORCID iD

Hermann Barbe  <https://orcid.org/0000-0003-1899-0981>

NOTES

1. An avatar is usually considered a digital representation of a person that is controlled by a real human over a computer interface.
2. The vignettes were short case descriptions of the virtual character, covering their social, family, and everyday situation so that trainees have some reference points for the conversation.
3. WER is the standard measurement to assess the quality of so-called automatic speech recognition systems (Morris et al., 2004). It is defined as the proportion of transcription errors to words processed in total. Errors are the number of wrongful substitutions, deletions or insertions of words.

REFERENCES

- Barbe, H., Siegel, B., Müller, J. L., & Fromberger, P. (2020). Welches Potenzial haben virtuelle Realitäten in der klinischen und forensischen Psychiatrie? Ein Überblick über aktuelle Verfahren und Einsatzmöglichkeiten [What is the potential of virtual reality in clinical and forensic psychiatry? An overview of current procedures and possible applications]. *Forensische Psychiatrie, Psychologie, Kriminologie*, 14(3), 270–277. <https://doi.org/10.1007/s11757-020-00611-2>
- Bohil, C. J., Alicea, B., & Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nature Reviews Neuroscience*, 12(12), 752–762. <https://doi.org/10.1038/nrn3122>
- Bordnick, P. S., & Washburn, M. (2019). Virtual environments for substance abuse assessment and treatment. In A. Rizzo & S. Bouchard (Eds.), *Virtual reality for psychological and neurocognitive interventions. Virtual reality technologies for health and clinical applications* (pp. 131–161). Springer. https://doi.org/10.1007/978-1-4939-9482-3_6
- Emmelkamp, P. M. G., & Meyerbröker, K. (2021). Virtual reality therapy in mental health. *Annual Review of Clinical Psychology*, 17(1), 1–25. <https://doi.org/10.1146/annurev-clinpsy-081219-115923>
- Fromberger, P., Jordan, K., & Müller, J. L. (2014). Anwendung virtueller Realitäten in der forensischen Psychiatrie [Application of virtual realities in forensic psychiatry]. *Der Nervenarzt*, 85(3), 298–303. <https://doi.org/10.1007/s00115-013-3904-7>
- Fromberger, P., Jordan, K., & Müller, J. L. (2018). Virtual reality applications for diagnosis, risk assessment and therapy of child abusers. *Behavioral Sciences & the Law*, 36(2), 235–244. <https://doi.org/10.1002/bsl.2332>
- Heyselaar, E., Hagoort, P., & Segaert, K. (2017). In dialogue with an avatar, language behavior is identical to dialogue with a human partner. *Behavior Research Methods*, 49(1), 46–60. <https://doi.org/10.3758/s13428-015-0688-7>
- Johnsen, K., Dickerson, R., Raij, A., Lok, B., Jackson, J., Shin, M., Hernandez, J., Stevens, A., & Lind, D. S. (2005). *Experiences in using immersive virtual characters to educate medical communication skills* [Conference session]. IEEE Proceedings. VR 2005. Virtual Reality, 2005, Bonn, Germany, pp. 179–186. <https://doi.org/10.1109/VR.2005.1492772>
- Kaplan, A. D., Cruit, J., Endsley, M., Beers, S. M., Sawyer, B. D., & Hancock, P. A. (2020). The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: A meta-analysis. *Human Factors: The Journal of Human Factors and Ergonomics Society*, 63(4), 706–726. <https://doi.org/10.1177/0018720820904229>
- Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., & Mitchell, S. (2010). Is ongoing feedback necessary to maintain the quality of investigative interviews with allegedly abused children? *Applied Developmental Science*, 6(1), 35–41. https://doi.org/10.1207/s1532480xads0601_04
- Maicher, K., Danforth, D., Price, A., Zimmerman, L., Wilcox, B., Liston, B., Cronau, H., Belknap, L., Ledford, C., Way, D., Post, D., Macerollo, A., & Rizer, M. (2017). Developing a conversational virtual standardized patient to enable students to practice history-taking skills. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, 12(2), 124–131. <https://doi.org/10.1097/sih.0000000000000195>
- Maicher, K., Zimmerman, L., Wilcox, B., Liston, B., Cronau, H., Macerollo, A., Jin, L., Jaffe, E., White, M., Fosler-Lussier, E., Schuler, W., Way, D. P., & Danforth, D. R. (2019). Using virtual standardized patients to accurately assess information gathering skills in medical students. *Medical Teacher*, 41(9), 1053–1059. <https://doi.org/10.1080/0142159x.2019.1616683>
- Milde, B., & Köhn, A. (2018). *Open source automatic speech recognition for German* [Symposium]. *Speech Communication; 13th ITG-Symposium 2018, VDE, Frankfurt am Main*, pp. 1–5.
- Morris, A. C., Maier, V., & Green, P. D. (2004). *From WER and RIL to MER and WIL: Improved evaluation measures for connected speech recognition*. https://www.researchgate.net/publication/221478089_From_WER_and_RIL_to_MER_and_WIL_improved_evaluation_measures_for_connected_speech_recognition
- Parsons, T. D., & Rizzo, A. “Skip.” (2019). A review of virtual classroom environments for neuropsychological assessment. In A. Rizzo & S. Bouchard (Eds.), *Virtual reality for psychological and neurocognitive interventions. Virtual reality technologies for health and clinical applications* (pp. 247–265). Springer. https://doi.org/10.1007/978-1-4939-9482-3_11

- Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., Hannemann, M., Motlicek, P., Qian, Y., Schwarz, P., Silovsky, J., Stemmer, G., & Vesely, K. (2011). *The Kaldi speech recognition toolkit* [Workshop]. *IEEE 2011 Workshop on Automatic Speech Recognition and Understanding, Hilton Waikoloa Village, Big Island, Hawaii, United State*.
- Ryan, W. S., Cornick, J., Blascovich, J., & Bailenson, J. N. (2019). Virtual reality: Whence, how and what for. In A. Rizzo & S. Bouchard (Eds.), *Virtual reality for psychological and neurocognitive interventions. Virtual reality technologies for health and clinical applications* (pp. 15–46). Springer. https://doi.org/10.1007/978-1-4939-9482-3_2
- Sygel, K., & Wallinius, M. (2021). Immersive virtual reality simulation in forensic psychiatry and adjacent clinical fields: A review of current assessment and treatment methods for practitioners. *Frontiers in Psychiatry, 12*, 673089. <https://doi.org/10.3389/fpsy.2021.673089>
- Talbot, T., & Rizzo, A. "Skip." (2019). Virtual human standardized patients for clinical training. In A. Rizzo & S. Bouchard (Eds.), *Virtual reality for psychological and neurocognitive interventions. Virtual reality technologies for health and clinical applications* (pp. 387–405). Springer. https://doi.org/10.1007/978-1-4939-9482-3_17
- Talbot, T., Sagae, K., John, B., & Rizzo, A. A. (2012). Sorting out the virtual patient: How to exploit artificial intelligence, game technology and sound educational practices to create engaging role-playing simulations. *International Journal of Gaming and Computer-Mediated Simulations (IJGMS), 4*(3), 1–19. <https://doi.org/10.4018/jgms.2012070101>
- Wang, R., DeMaria, S., Goldberg, A., & Katz, D. (2016). A systematic review of serious games in training health care professionals. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare, 11*(1), 41–51. <https://doi.org/10.1097/sih.0000000000000118>
- Wilcox, B., & Wilcox, S. (2013). Making it real: Loebner-winning chatbot design. *Arbor, 189*(764), a086. <https://doi.org/10.3989/arbor.2013.764n6009>

Hermann Barbe has a master's degree in psychology and currently works as a research assistant at the Human Medical Center Göttingen, Clinic for Psychiatry and Psychotherapy—Forensic Psychiatry, Germany. His research involves the development of new research tools using digital technologies (Web Applications, Virtual Reality) and their use in the field of forensic psychiatry and forensic psychotherapy

Bruno Siegel has a master's degree in psychology and is a licensed psychotherapist. His research interests include the use of new media in forensic psychiatry and forensic psychotherapy.

Jürgen L. Müller, MD, is a professor of Forensic Psychiatry at the Georg August University Göttingen and head of the Asklepios Clinic for Forensic Psychiatry.

Peter Fromberger, PhD, is a postdoctoral researcher at the Human Medical Center Göttingen, Clinic for Psychiatry and Psychotherapy—Forensic Psychiatry, Germany. Peter's doctoral thesis examined the possibilities to detect deviant pedosexual interests based on eye movements. Peter received the Eberhard Schorsch Award of the German Association for Sexual Research (DGfS) and the award for Empirical Research in Forensic Psychiatry of the German Association for Psychiatry and Psychotherapy (DGPPN) for his thesis. His post-doc research comprises the usability of digital techniques (Virtual Reality, web-based interventions) in the context of sexual deviancy.