Students’ attitudes towards AI in Psychiatry

Anne, Susanne, & Eesha

# Seed for random number generation  
set.seed(42)

# Introduction

Over the past decades, little has changed concerning how psychotherapy is delivered (Johnsen & Friborg, 2015). A lack of targeted feedback, care quality monitoring, and stagnation in terms of further education and training of (prospective) psychotherapists may hinder the implementation of effective interventions (Cummins et al., 2019; Hirsch et al., 2018; Schwalbe et al., 2014). Artificial intelligence (AI) technologies for mental health care may address these problems. AI-enabled tools may improve the quality of psychotherapeutic training and education. To this end, AI systems analyze data gathered from therapist-patient conversations to provide performance-specific feedback for the therapist, thus potentially enhancing their motivational interviewing performance (Cummins et al., 2019; Hirsch et al., 2018; Imel et al., 2019; Tanana et al., 2019a).

Despite a large amount of academic knowledge and efforts to develop user-friendly applications, AI systems are still hardly utilized in clinical care (Sendak et al., 2020). Technical and administrative difficulties, such as the inaccuracy of predictions, accuracy-interpretability trade-offs, and data privacy and security concerns complicate the widespread implementation of AI tools in diagnosis, prognosis, and the selection of treatment approaches (Aafjes-van Doorn et al., 2021; Chekroud et al., 2021; Chen et al., 2022; Kelly et al., 2019; Lee et al., 2021; Roth et al., 2021). Next to general reasons for reluctance against using AI-enabled tools in mental healthcare, specific applications may be associated with specific forms of skepticism. Regarding AI-enabled training devices, based on the belief that an algorithm may not judge human-provided care, psychotherapists may be hesitant to accept feedback regarding their therapeutic practice from a machine. The skepticism against AI-enabled tools in mental healthcare is nurtured by a lack of understanding of how AI recommendations are generated (Aafjes-van Doorn et al., 2021; Chekroud et al., 2021). Despite attempts to enhance the explainability of AI, such as the Explainable Artificial Intelligence (XAI) Initiative, the complexity of deep learning approaches necessarily limits the extent to which they can be made accessible to a broader user group (Feldman et al., 2019). Especially in mental healthcare, where transparency and the explainability of clinical decision-making are highly valued, the black box problem of AI-based recommendations creates a significant obstacle to its adoption (Aafjes-van Doorn et al., 2021; Chekroud et al., 2021; Kelly et al., 2019).

## AI-enabled feedback tools in psychotherapy training

Supervision and receiving performance feedback on their therapy sessions support psychotherapy trainees’ skills acquisition and increase retention Moyers et al. (2005). However, providing ongoing feedback is labor and cost intensive and thus rarely used in training and clinical practice. Accordingly, feedback is often based on trainees’ self-reports and is usually only available long after the session (Tanana et al., 2019b). Using AI technology for training purposes in mental health care may lhelp to reduce this problem by providing continuous, immediate, and performance-specific feedback to psychotherapists and trainees.

Most tools developed to improve psychotherapy quality rely on natural language processing-based feedback Tanana et al. (2019b). For example, *TIM* (Therapy Insights Model) uses real-time chat messages exchanged between therapists and patients to provide therapists with feedback regarding the topics that were sufficiently covered during the session and the topics that should be addressed in the following sessions (Cummins et al., 2019). *CORE-MI* (Counselor Observer Ratings Expert for Motivational Interviewing) uses audio recordings of motivational interviewing (MI)[[1]](#footnote-1) sessions to generate feedback on psychotherapists’ adherence to MI principles. The user receives feedback on six summary measures of MI fidelity: empathy, MI spirit, reflection-to-question ratio, percent open questions, percent complex reflections, and percent MI adherence. *CORE-MI* includes a visual summary of counseling sessions based on the fidelity assessment that the therapist may use to improve their MI performance (Hirsch et al., 2018). Similar tools include the *ClientBot*, a training tool that mimics typical patient responses to therapist questions and provides real-time feedback on therapists’ use of open questions and reflections (Tanana et al., 2019b); or *Partner*, a reinforcement learning agent that may increase the quality of mental health support conversations by suggesting sentence-level edits to posts that enhance the level of empathy while maintaining conversation quality (Sharma et al., 2021). The feedback tool we use in the current study analyses therapeutic conversations between practitioner and patient to deliver targeted feedback to psychotherapists based on the principles of motivational interviewing [see the Therapy Insights Model (TIM); Cummins et al. (2019)].

## The Unified Theory of Acceptance and Use of Technology

The unified theory of acceptance and use of technology [UTAUT; Venkatesh (2022), Venkatesh et al. (2003), Venkatesh et al. (2016)] provides a theoretical framework that explains the relationship between technology, environment, and user characteristics with the behavioral intention to use an AI-enabled tool. The UTAUT includes four main predictors of the intention to use a tool: a) performance expectancy, defined as the degree to which an individual believes that using a system will enhance their performance, b) effort expectancy, as the degree of ease associated with using the technology, c) social influence, referring to the perception that important others believe that the system should be used, and d) facilitating conditions, as the belief that the infrastructure exists to support the use of the system. Over the past decades, modified versions of the UTAUT have been developed and applied to study the acceptance and intention to use AI-enabled tools in healthcare [arfi\_etal21; gado\_etal22a; Fan et al. (2020); Tran et al. (2021); Zhai et al. (2021); Tamori et al. (2022)]. Because of the high-stakes decision-making process and the sensitive nature of the data used for AI-enabled recommendation systems in healthcare, perceived trust has been proposed to act as a relevant additional predictor of the behavioral intention to use AI-enabled tools in healthcare (Arfi et al., 2021).

HERE STH ON WHY UTAUT AND NOT OTHER THEORIES:

* because social influence is included and may be relevant among students
* we address the criticism of the UTAUT (Shachak et al., 2019):

1. Value adding use is often neglected in UTAUT research (Shachak et al., 2019); we do not focus on how the AI tool is used in practice. BUT: the prerequisite for value adding use is acceptance and openness towards a technology; by investigating the hypotheses in a student sample and focusing on the intention to use the tool in their future jobs, we shed light on the processes that enable value adding use
2. Adopt and develop theoretical frameworks and methodologies that account for multiple, interrelated, sociotechnical aspects (Shachak et al., 2019): Because students are not yet operating in an organizational context, influencing factors are limited to their social contexts and the educational setting; accordingly, our research allows a stronger focus on the individual predictors of the intention to use the tool; without confounding by organizational setting, work tasks, and habits
3. same applied to “Accounting for health system complexity” (Shachak et al., 2019)
4. Understanding and reconciling multiple user needs (Shachak et al., 2019): We selected a relatively homogenous sample of psychology master’s students - reconciliation of multiple user needs does not play a major role
5. Consider temporal dimensions of HIT implementation (Shachak et al., 2019): We do not focus on an implementation setting, but on general openness towards using the tool.

# Hypotheses Development

Based on previous research findings, the first goal of the current research is to test the applicability of a modified version of the UTAUT in the mental health context to understand the factors that influence the willingness to accept AI-enabled recommendations (Gado et al., 2022; Venkatesh, 2022; Venkatesh et al., 2016; Venkatesh et al., 2003). Specifically, in Study 1, we investigate the relevance of UTAUT predictors for the intention to use an AI-enabled feedback tool among samples of psychology students specialized in clinical psychology. In line with the UTAUT, we propose performance expectancy and effort expectancy to predict the behavioral intention to use the feedback tool. That is, students who perceive the tool as useful and believe that it would be easy for them to use the tool likely score high on the behavioral intention to use it.

Hypothesis 1: There is a positive relationship between psychology students’ perceived performance expectancy of the tool and their intention to use the tool in their future job.

Hypothesis 2: There is a positive relationship between psychology students’ perceived effort expectancy and their intention to use the tool in their future job.

In contrast to practicing psychotherapists, psychology students are less influenced by habits and established work processes that may hinder the adoption of new AI technologies (Venkatesh et al., 2016). Accordingly, students may be susceptible to the influence of their peers and the perceived norms and values of their future employers. Thus, we suggest the UTUAT variable social influence as a predictor of students’ intention to use the feedback tool.

Hypothesis 3: There is a positive relationship between social influence and students’ intention to use the tool in their future job.

Due to the sensitive nature of the data used and the recommendations made by the tool, we propose trust to act as a predictor of students’ intention to use the feedback tool (Arfi et al., 2021).

Hypothesis 4: There is a positive relationship between students’ trust in the tool and their intention to use the tool in their future job.

To our knowledge, only one study has investigated the predictors of technology acceptance and use in the mental health domain [gado\_etal22]. Gado et al. (2022) found that general perceived usefulness and perceived ease of use positively predicted intention to use AI tools. The effect was mediated through favorable attitudes towards AI. In addition, the authors identified knowledge about AI tools as a direct predictor of the intention to use the tool. We argue that the relationship between knowledge and intention to use is mediated through the UTAUT variables performance expectancy, effort expectancy, and trust. The idea that knowledge predicts performance expectancy has been put forward by other theories of technology acceptance and use, such as the New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies (NASSS, Greenhalgh et al., 2017). According to the NASSS framework, knowledge of a tool predicts its perceived demand-side value. That is, students who possess the knowledge and skills necessary to apply the tool and understand how the recommendations are derived are more likely to perceive it as useful. Several training and intervention studies have shown that adequate education efforts with the goal of building knowledge impact on clinicians’ views on adoption and their ongoing use of the system (Bredfeldt et al., 2013; Kraus et al., 2008; Wiljer et al., 2021). Finally, an understanding of how the AI recommendations are derived may strengthen students’ competence in using the tool in their future jobs and leverage some ethical concerns, thus potentially increasing their effort expectancy and trust in the tool (Gado et al., 2022; Seufert et al., 2021). Accordingly, we extend the UTAUT model by including technology readiness as an indicator of general AI knowledge and understanding of the tool as an indicator of specific AI knowledge as predictors of performance expectancy, effort expectancy, and trust.

Hypothesis 5: The positive relationship between technology readiness and the intention to use the tool is mediated through a) performance expectancy, b) effort expectancy, and c) trust in the tool.

Hypothesis 6: The positive relationship between understanding of the tool and the intention to use the tool is mediated through a) performance expectancy, b) effort expectancy, and c) trust in the tool.

The second goal of the current research is to examine the effectiveness of a skill-based intervention on the intention to use the two previously described AI tools. Thus, based on the findings of Study 1, in Study 2, we test the effects of a skill-based intervention on students’ intention to use the two AI tools described above.

# Study 1

## Methods

### Participants

(**pintodossantos\_etal19?**): “The questionnaire was sent out via email and advertised on social media to undergraduate medical students at three major German universities (University of Cologne, University of Bonn, University of Mainz). Participation was voluntary and had no relation to the student’s curricular activities. The students were informed that the survey results would be used for further statistical evaluation and scientific publication. Respondent anonymity was guaranteed by design […].”

### Measurement Instruments

#### AI tools

##### Therapist Feedback Tool

|  |
| --- |
| Core MI Feedback Tool |

#### Independent Variables Based on UTAUT

* Perceived social norm Gado et al. (2022); all slightly adapted:
* Performance expectancy (perceived usefulness) Gado et al. (2022); all slightly adapted:
* Effort expectancy (perceived ease of use) Gado et al. (2022); all slightly adapted:
* Knowledge of the tool (Gado et al., 2022):
  + “Please rate your understanding of how the recommendations delivered by [short tool description] are derived [in comparison to your fellow students].”

#### Dependent Variable: Intention to Use the Tool

* Intention to use the tool Gado et al. (2022); all slightly adapted:

#### Additional Variables and Control Variables

* Technology readiness
* Perceived trust in the tool/ credibility
* Professional identity
* General technology affinity
* Computer self-efficacy
* Technostress
* Relevant education content (stats course)
* Personality (Park & Woo, 2022)
* Data privacy concerns
* affective, cognitive, behavioral attitudes towards AI (Park & Woo, 2022)

# References

Aafjes-van Doorn, K., Kamsteeg, C., Bate, J., & Aafjes, M. (2021). A scoping review of machine learning in psychotherapy research. *Psychotherapy Research*, *31*(1), 92–116. <https://doi.org/10.1080/10503307.2020.1808729>

Arfi, W. B., Nasr, I. B., Kondrateva, G., & Hikkerova, L. (2021). The role of trust in intention to use the IoT in eHealth: Application of the modified UTAUT in a consumer context. *Technological Forecasting and Social Change*, *167*, 120688. <https://doi.org/10.1016/j.techfore.2021.120688>

Atkins, D. C., Steyvers, M., Imel, Z. E., & Smyth, P. (2014). Scaling up the evaluation of psychotherapy: Evaluating motivational interviewing fidelity via statistical text classification. *Implementation Science: IS*, *9*, 49. <https://doi.org/10.1186/1748-5908-9-49>

Bredfeldt, C. E., Awad, E. B., Joseph, K., & Snyder, M. H. (2013). Training providers: Beyond the basics of electronic health records. *BMC Health Services Research*, *13*(1), 503. <https://doi.org/10.1186/1472-6963-13-503>

Can, D., Marín, R. A., Georgiou, P. G., Imel, Z. E., Atkins, D. C., & Narayanan, S. S. (2016). "It sounds like...": A natural language processing approach to detecting counselor reflections in motivational interviewing. *Journal of Counseling Psychology*, *63*(3), 343–350. <https://doi.org/10.1037/cou0000111>

Chekroud, A. M., Bondar, J., Delgadillo, J., Doherty, G., Wasil, A., Fokkema, M., Cohen, Z., Belgrave, D., DeRubeis, R., Iniesta, R., Dwyer, D., & Choi, K. (2021). The promise of machine learning in predicting treatment outcomes in psychiatry. *World Psychiatry*, *20*(2), 154–170. <https://doi.org/10.1002/wps.20882>

Chen, Z. S., Prathamesh, Kulkarni, Galatzer-Levy, I. R., Bigio, B., Nasca, C., & Zhang, Y. (2022). *Modern Views of Machine Learning for Precision Psychiatry* (No. arXiv:2204.01607). arXiv. <https://arxiv.org/abs/2204.01607>

Cummins, R., Ewbank, M., Martin, A., Tablan, V., Catarino, A., & Blackwell, A. (2019). *TIM: A Tool for Gaining Insights into Psychotherapy* (p. 3506). <https://doi.org/10.1145/3308558.3314128>

Fan, W., Liu, J., Zhu, S., & Pardalos, P. M. (2020). Investigating the impacting factors for the healthcare professionals to adopt artificial intelligence-based medical diagnosis support system (AIMDSS). *Annals of Operations Research*, *294*(1), 567–592. <https://doi.org/10.1007/s10479-018-2818-y>

Feldman, R., Aldana, E., & Stein, K. (2019). Artificial Intelligence in the Health care Space: How We Can Trust What We Cannot Know. *Policy Review*, *30*, 23.

Gado, S., Kempen, R., Lingelbach, K., & Bipp, T. (2022). Artificial intelligence in psychology: How can we enable psychology students to accept and use artificial intelligence? *Psychology Learning & Teaching*, *21*(1), 37–56. <https://doi.org/10.1177/14757257211037149>

Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., A’Court, C., Hinder, S., Fahy, N., Procter, R., & Shaw, S. (2017). Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. *Journal of Medical Internet Research*, *19*(11), e8775. <https://doi.org/10.2196/jmir.8775>

Helge Rønnestad, M., & Ladany, N. (2006). The impact of psychotherapy training: Introduction to the special section. *Psychotherapy Research*, *16*(3), 261–267. <https://doi.org/10.1080/10503300600612241>

Hirsch, T., Soma, C., Merced, K., Kuo, P., Dembe, A., Caperton, D. D., Atkins, D. C., & Imel, Z. E. (2018). “It’s hard to argue with a computer:” Investigating Psychotherapists’ Attitudes towards Automated Evaluation. *DIS. Designing Interactive Systems (Conference)*, *2018*, 559–571. <https://doi.org/10.1145/3196709.3196776>

Imel, Z. E., Pace, B. T., Soma, C. S., Tanana, M., Hirsch, T., Gibson, J., Georgiou, P., Narayanan, S., & Atkins, D. C. (2019). Design feasibility of an automated, machine-learning based feedback system for motivational interviewing. *Psychotherapy*, *56*(2), 318–328. <https://doi.org/10.1037/pst0000221>

Johnsen, T. J., & Friborg, O. (2015). The effects of cognitive behavioral therapy as an anti-depressive treatment is falling: A meta-analysis. *Psychological Bulletin*, *141*(4), 747–768. <https://doi.org/10.1037/bul0000015>

Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, *17*(1), 195. <https://doi.org/10.1186/s12916-019-1426-2>

Kraus, S., Barber, T. R., Briggs, B., Kall, G., Moran, B., Vernon, T. R., & Waggoner, D. M. (2008). Implementing Computerized Physician Order Management at a Community Hospital. *The Joint Commission Journal on Quality and Patient Safety*, *34*(2), 74–84. <https://doi.org/10.1016/S1553-7250(08)34010-0>

Lee, E. E., Torous, J., De Choudhury, M., Depp, C. A., Graham, S. A., Kim, H.-C., Paulus, M. P., Krystal, J. H., & Jeste, D. V. (2021). Artificial Intelligence for Mental Health Care: Clinical Applications, Barriers, Facilitators, and Artificial Wisdom. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *6*(9), 856–864. <https://doi.org/10.1016/j.bpsc.2021.02.001>

Moyers, T. B., Martin, T., Manuel, J. K., Hendrickson, S. M. L., & Miller, W. R. (2005). Assessing competence in the use of motivational interviewing. *Journal of Substance Abuse Treatment*, *28*(1), 19–26. <https://doi.org/10.1016/j.jsat.2004.11.001>

Park, J., & Woo, S. E. (2022). Who Likes Artificial Intelligence? Personality Predictors of Attitudes toward Artificial Intelligence. *The Journal of Psychology*, *156*(1), 68–94. <https://doi.org/10.1080/00223980.2021.2012109>

Roth, C. B., Papassotiropoulos, A., Brühl, A. B., Lang, U. E., & Huber, C. G. (2021). Psychiatry in the Digital Age: A Blessing or a Curse? *International Journal of Environmental Research and Public Health*, *18*(16), 8302. <https://doi.org/10.3390/ijerph18168302>

Schwalbe, C. S., Oh, H. Y., & Zweben, A. (2014). Sustaining motivational interviewing: A meta-analysis of training studies. *Addiction*, *109*(8), 1287–1294. <https://doi.org/10.1111/add.12558>

Sendak, M. P., D’Arcy, J., Kashyap, S., Gao, M., Nichols, M., Corey, K., Ratliff, W., & Balu, S. (2020). A Path for Translation of Machine Learning Products into Healthcare Delivery. *EMJ Innovations*. <https://doi.org/10.33590/emjinnov/19-00172>

Seufert, S., Guggemos, J., & Sailer, M. (2021). Technology-related knowledge, skills, and attitudes of pre- and in-service teachers: The current situation and emerging trends. *Computers in Human Behavior*, *115*, 106552. <https://doi.org/10.1016/j.chb.2020.106552>

Shachak, A., Kuziemsky, C., & Petersen, C. (2019). Beyond TAM and UTAUT: Future directions for HIT implementation research. *Journal of Biomedical Informatics*, *100*, 103315. <https://doi.org/10.1016/j.jbi.2019.103315>

Sharma, A., Lin, I. W., Miner, A. S., Atkins, D. C., & Althoff, T. (2021). Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. *Proceedings of the Web Conference 2021*, 194–205. <https://doi.org/10.1145/3442381.3450097>

Tamori, H., Yamashina, H., Mukai, M., Morii, Y., Suzuki, T., & Ogasawara, K. (2022). Acceptance of the Use of Artificial Intelligence in Medicine Among Japan’s Doctors and the Public: A Questionnaire Survey. *JMIR Human Factors*, *9*(1), e24680. <https://doi.org/10.2196/24680>

Tanana, M. J., Soma, C. S., Srikumar, V., Atkins, D. C., & Imel, Z. E. (2019b). Development and Evaluation of ClientBot: Patient-Like Conversational Agent to Train Basic Counseling Skills. *Journal of Medical Internet Research*, *21*(7), e12529. <https://doi.org/10.2196/12529>

Tanana, M. J., Soma, C. S., Srikumar, V., Atkins, D. C., & Imel, Z. E. (2019a). Development and Evaluation of ClientBot: Patient-Like Conversational Agent to Train Basic Counseling Skills. *Journal of Medical Internet Research*, *21*(7), e12529. <https://doi.org/10.2196/12529>

Tran, A. Q., Nguyen, L. H., Nguyen, H. S. A., Nguyen, C. T., Vu, L. G., Zhang, M., Vu, T. M. T., Nguyen, S. H., Tran, B. X., Latkin, C. A., Ho, R. C. M., & Ho, C. S. H. (2021). Determinants of Intention to Use Artificial Intelligence-Based Diagnosis Support System Among Prospective Physicians. *Frontiers in Public Health*, *9*, 755644. <https://doi.org/10.3389/fpubh.2021.755644>

Venkatesh, V. (2022). Adoption and use of AI tools: A research agenda grounded in UTAUT. *Annals of Operations Research*, *308*(1-2), 641–652. <https://doi.org/10.1007/s10479-020-03918-9>

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425–478. <https://doi.org/10.2307/30036540>

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2016). *Unified Theory of Acceptance and Use of Technology: A Synthesis and the Road Ahead* ({{SSRN Scholarly Paper}} No. 2800121).

Wiljer, D., Salhia, M., Dolatabadi, E., Dhalla, A., Gillan, C., Al-Mouaswas, D., Jackson, E., Waldorf, J., Mattson, J., Clare, M., Lalani, N., Charow, R., Balakumar, S., Younus, S., Jeyakumar, T., Peteanu, W., & Tavares, W. (2021). Accelerating the Appropriate Adoption of Artificial Intelligence in Health Care: Protocol for a Multistepped Approach. *JMIR Research Protocols*, *10*(10), e30940. <https://doi.org/10.2196/30940>

Zhai, H., Yang, X., Xue, J., Lavender, C., Ye, T., Li, J.-B., Xu, L., Lin, L., Cao, W., & Sun, Y. (2021). Radiation Oncologists’ Perceptions of Adopting an Artificial Intelligence–Assisted Contouring Technology: Model Development and Questionnaire Study. *Journal of Medical Internet Research*, *23*(9), e27122. <https://doi.org/10.2196/27122>

1. A counseling method to enhance a patient’s motivation to change. [↑](#footnote-ref-1)