Journal of Technology in Behavioral Science
https://doi.org/10.1007/s41347-025-00484-4

Teaming with Artificial Intelligence to Learn and Sustain Psychotherapy Delivery Skills: Workplace, Ethical, and Research Implications

Andrew M. Sherrill[1] · Christopher W. Wiese[2] · Saeed Abdullah[3] · Rosa I. Arriaga[2]

Received: 10 May 2024 / Revised: 4 November 2024 / Accepted: 20 January 2025

Abstract
The shortage of psychotherapists is anticipated to escalate over the next few decades, necessitating interdisciplinary responses to create scalable solutions to meet workforce demands. This paper explores the potential of teaming with artificial intelligence (AI) to upskill current and future psychotherapists. Specifically, we discuss the benefits of AI teaming in augmenting human clinicians' capacity to learn empirically supported treatments (ESTs), thus improving accessibility to the best available treatments. We argue that integrating AI as a teammate rather than a tool can facilitate the ethical and effective technology integration. While AI interfaces may create the illusion of teamwork between humans and synthetic agents, we also suggest that an "AI teammate" not be endowed with a human teammate's capacities to be held personally accountable and understand the roles and responsibilities of other teammates. We highlight the necessity of addressing workforce implications, such as changes in collaborative dynamics and competency requirements. Ethical considerations, including transparency, fairness, and privacy, must also guide the integration of AI into mental health work to prevent unintended biases and ensure responsible use. Additionally, we discuss research implications, emphasizing the need to move beyond understanding how psychotherapists interact with AI to how they will collaborate with it. Finally, we outline technological innovations needed for successful AI teaming, including personalized feedback, usability, and bidirectional communication structures. Teaming with AI has the potential to transform the mental health workforce, but collaboration across stakeholders and adherence to ethical principles are essential for successful integration.

Keywords Psychotherapy · Training · Dissemination · Artificial intelligence · Human-AI teams

Over the next several decades, the shortage of psychotherapists in the United States is expected to reach the hundreds of thousands (Health Resources and Services Administration, 2024). Humans alone may not address the demand for psychotherapy. An interdisciplinary response between clinical scientists, information scientists, and organizational scientists is needed to create scalable solutions to make effective treatment broadly accessible (Koerner et al., 2022). In this paper, we address the potential benefits afforded by teaming with artificial intelligence (AI) in mental health work with the specific use cases of learning and sustaining psychotherapy delivery skills. We describe how AI may potentially serve some functions of a supervisor for novice psychotherapists and some functions of a consultant for established psychotherapists who want to improve competencies or prevent gradual drift from best practices (e.g., Speers et al., 2022). Next, we highlight workforce, ethical, and research implications that will emerge with AI teaming in mental health work. Lastly, we discuss the technological innovations required for AI teaming to be useful and usable.

✉ Andrew M. Sherrill
andrew.m.sherrill@emory.edu

[1] Emory University School of Medicine, 12 Executive Park Drive NE, 3Rd Floor, Atlanta, GA 30326, USA

[2] Georgia Institute of Technology, Atlanta, GA, USA

[3] Pennsyvania State University, University Park, PA, USA

Potential Benefits of AI Teaming

Broadly speaking, essential benefits of AI are to streamline, augment, and supplement human work. Discussion of all potential aspects of mental health work that may benefit from AI is beyond the scope of this paper (e.g., identifying of at-risk individuals, interpreting assessments, and delivering psychotherapy). We draw attention to the potential of AI teaming within the domain of human clinicians learning and sustaining psychotherapy delivery skills. We define AI teaming as collaboration between artificial and human agents for the purpose of working toward a shared goal (Berretta et al., 2023). While AI interfaces may create the illusion of an "AI teammate," AI should not be confused with human teammates who can understand the roles and responsibilities of other teammates and be held personally accountable. AI teammates merely extend the capabilities of licensed practitioners and therefore cannot be held accountable. Clarifying the distinction between human teammates and AI teammates is essential to understand the workforce, ethical, and research implications of "AI teaming" in mental health work.

Importantly, many of the most effective psychotherapy protocols that are supported by clinical trials, known as empirically supported treatments (ESTs), are accessible only to a small minority of mental health patients (Becker et al., 2013; Cho et al., 2019; Kessler et al., 2005). Successful implementation of ESTs depends on many factors (Damschroder et al., 2015; Gallo & Barlow, 2012); however, research indicates one of the biggest barriers to upskilling the workforce is the availability of training (Brennan et al., 2022; Cook et al., 2009a), as opportunities to actively learn new skills predict the implementation of new skills Cook et al., 2009b; Henrich et al., 2023; Karlin & Cross, 2014).

We believe teaming with AI can revolutionize the availability and effectiveness of EST training by reducing the reliance on passive didactic experiences (Beidas et al., 2012; Henrich et al., 2023; Valenstein-Mah et al., 2020) and resource-consuming individual consultation (Bash & Stirman, 2020). EST experts who provide didactics and consultation represent a relatively minuscule subgroup of psychotherapists and, therefore, they cannot fully support growing workforce demands (Amerikaner & Rose, 2012; Frank et al., 2020). Due to the resource-prohibitive nature of teaming with experts, psychotherapists are limited in their ability to learn and implement EST protocols.

Teams are defined as two or more autonomous entities that work interdependently to achieve a shared goal (Salas et al., 1992). When feasible, teaming with experts works in EST training (e.g., Sherrill et al., 2021); however, its accessibility is limited (Bash & Stirman, 2020). Teaming may continue to be the best solution, though perhaps with the integration of AI into the team. Importantly, AI systems are developing to such a point where the technology is experienced by the user as a *teammate* (O'Neill et al., 2022). The integration of an *AI teammate* has far greater implications for the future of mental health work than simply using AI as a data processing tool. For instance, an AI teammate might not only passively observe and provide feedback on therapy sessions but also receive corrective feedback from the psychotherapist, thus enabling shared cognition within the worker-AI team. Ideally, these bidirectional AI systems will function as objective, nonjudgmental, automatic, ever-present, and confidential consultants who can provide individualized feedback throughout the psychotherapist's career. This idea of teaming with AI goes beyond how they are currently implemented in the mental health workforce. While there has been considerable progress in developing consumer-grade AI tools to enhance fidelity (e.g., Lyssn; www.lyssn.io), these tools provide unidirectional *coaching* to improve the psychotherapist's treatment fidelity (i.e., consistency with protocol), where the psychotherapist receives feedback to help them to reflect on their practice (Creed et al., 2022; Flemotomos et al., 2021; Xiao et al., 2015). Future computational systems will transform the mental health workforce by transitioning from unidirectional AI tools to bidirectional worker-AI teams (O'Neill et al., 2022). These synthetic agents can provide feedback, receive feedback, and dynamically update their models. Research is needed to understand not only how to build and design these bidirectional systems for this work context but also how to integrate them ethically and effectively into diverse clinical settings.

Implications of AI Teaming

Workforce Implications

While there are broad workforce implications to incorporating AI, there are implications specific to integrating AI teammates into the process of learning and sustaining psychotherapy skills. Specifically, there is a gap in the research regarding end-user preferences and expectations of AI teammates within the contexts of learning and delivering psychotherapy. Prior research within the context of multiplayer gaming found that instrumental skills were end-users' top priority in an AI teammate (Zhang et al., 2021); however, little research has explored the extent to which preferences and expectations in human-AI teaming generalize across contexts. Further, research is needed to understand how psychotherapists' requirements fit into the requirements of other stakeholders (e.g., patients and administrators) and how to design AI systems that all stakeholders perceive as ethical, useful, and useable. Stakeholders are likely to evaluate and interact with AI teammates fundamentally differently than their human teammates (Hidalgo et al., 2021).

The process of learning and sustaining effective EST delivery skills often leverages supervision or consultation from another psychotherapist with expertise in those skills. Supervision and consultation in EST is a complex interpersonal process (Terjesen & Del Vecchio, 2023) and little research indicates what are the most effective practices with regard to the learner's successful implementation of ESTs (Valenstein-Mah et al.., 2020; Zukerman et al., 2023). One supervision and consultation strategy is to improve the psychotherapist's adherence to a fidelity protocol by reviewing data collected during clinical sessions (e.g., video recordings) and then giving actionable feedback to the psychotherapist (e.g., Sherrill et al., 2021). These fidelity protocols are often the same protocols used in clinical trials that provide evidence of the given treatment's efficacy. This strategy requires the supervisor or consultant to rate the psychotherapist's performance on established checklists of essential components of a given session (e.g., Burton et al., 2023). Additionally, the supervisor or consultant reflects on performance with subjective reflections (e.g., "Was the psychotherapist empathic?") that are based on tacit and ambiguous rules for what is appropriate for a given psychotherapist at a given moment within a given protocol. Importantly, there are situational (e.g., supervisor availability), practical (e.g., memory-related errors of sessions), and perceptual (e.g., biased recollections of effectiveness) factors that prevent humans from effectively monitoring performance and making necessary improvements. If an AI teammate were including the work task of determining effective protocol delivery, its input will likely use objective and computational sensors (e.g., acoustic analysis of the therapist's voice, natural language processing of the therapist's work choices). Figure 1 depicts a hypothetical information flow within these future teams, illustrating how the complexity of this work task increases with the introduction of an AI teammate.

Initial integration will result in "growing pains" by increasing cognitive demands and stress levels of the human operators (Rosero et al., 2021; Woods et al., 2002). Further, AI integration may be prone to fail in instances of feedback and learning if social dynamics are not considered (e.g., the capacity of administration in critically viewing feedback from the AI teammate; Huang et al., 2021; McNeese et al., 2021). This is highlighted by the observations that psychotherapists may be apprehensive about adopting new technologies into their complicated work ecosystems (Lattie et al., 2020). For example, despite over a decade of empirical support, most psychotherapists did not use telehealth until they were required to during the COVID-19 pandemic (Dores et al., 2020; Sammons et al., 2020), and only then did they report positive attitudes toward that technology and intentions to continue integrating it into their work (Békés & Aafjes-van Doorn, 2020). To facilitate AI integration, risk mitigation frameworks will be needed (e.g., using exposure-based principles to overcome technophobia, Sherrill et al., 2022). Critical incident interviews can identify not only the ethical and practical challenges of introducing an AI teammate but also the socioemotional processes that led to the perception of these being "risks." Policy-capturing studies with subject matter experts can identify the efficacy of countermeasures used to mitigate the risks of integrating an AI teammate into mental health work.

The introduction of new technologies into mental health work will necessitate the re-examination of psychotherapist competencies necessary to perform the job well. The competencies needed to be a future psychotherapist are not necessarily those needed to succeed at this job currently. For one,

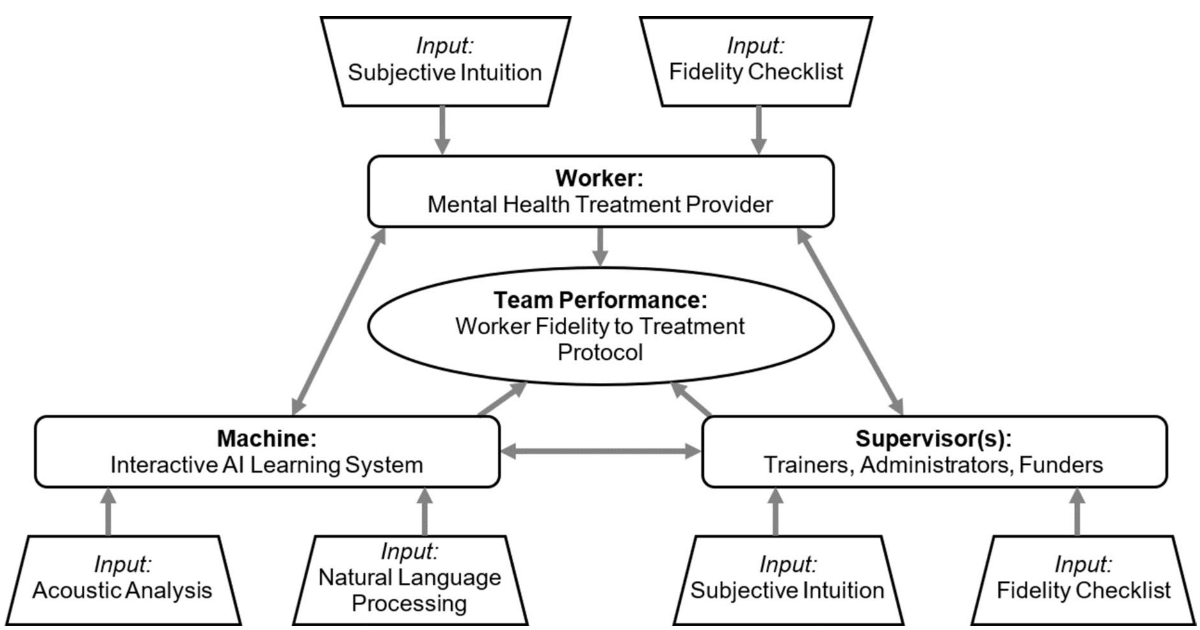

Fig. 1 Example worker-AI team information flow for a specific use case of learning treatment skills

the proficiency levels of many of the existing competencies will change, and research suggests new competencies will be needed to effectively team with AI (Demir et al., 2019; McNeese et al., 2018; Stowers et al., 2021). Communication, coordination, and adaptation will likely be essential teamwork competencies for worker-AI teams (Stowers et al., 2021). Hence, researchers need to (1) understand new taskwork dynamics in the future of mental health work and (2) identify the competencies of the future psychotherapist.

Central to understanding new workplace dynamics will be identifying the specific role of an AI teammate, which will be largely determined by characteristics of its human teammates. For example, when used as a tool in a private practice with a single licensed psychotherapist, the AI teammate will serve only the psychotherapist, perhaps by fulfilling the role of providing longitudinal feedback across patients on characteristics related to protocol fidelity, which the psychotherapist uses a protection against skill drift (Speers et al., 2022). In contrast, when used as a tool within the context of supervision, the AI teammate may primarily serve the supervisor, perhaps by fulfilling the roles of completing time-consuming tasks (e.g., scanning a 60-min therapy session recording for possible protocol deviations) and offering suggestions of discussions topics during supervision (e.g., closely reviewing a segment of a session recording that was flagged for a protocol deviation). The specific roles of the AI teammate within the team could impact each other teammate and relationships between them. For example, when used in the context of supervision, procedures will be needed to address disagreements between the AI teammate and supervisor, as will be determinations on how much access the supervisee might have to those disagreements. While learning how these AI tools are built and used is likely important, end-users will also benefit from articulating the roles of each agent (i.e., AI and humans) and anticipating how each human teammate and each human relationship may be affected by the introduction of an AI teammate.

Ethical Implications

The ethical implications of AI integration are crucial to consider in mental health work, as it chiefly serves vulnerable populations for which privacy and nonmaleficence are paramount. Broad ethical principles for AI applications (Jobin et al., 2019) include transparency (e.g., disclosure of approach, explainability of outputs), justice and fairness (e.g., prevention of bias, contestability of outputs), nonmaleficence (e.g., prevention of misuse, utilization of routine risk assessments), responsibility (e.g., clarifying liability, ensuring diversity in development), and privacy (e.g., protections for data security, obtaining full informed consent). Given the nascent stage, ethical reasoning in worker-AI teaming in mental health is speculative and demands targeted research (Fiske et al., 2019; Manriquez Roa et al., 2021). Chief among the challenges will be protecting clinical decision-making from unintended biases embedded in computational models. Integrating computational tools might introduce algorithmic biases that harm all stakeholders. Other challenges will be related to fundamental changes to the nature of work, including the ethical transfer of new responsibilities in AI operation (e.g., how to interpret feedback from AI teammates). The future of AI in mental health work must address how to ethically (1) collaborate with AI and (2) integrate transformational systems into the workplace. Without understanding the psychotherapist's readiness and needs, current AI technologies face unknown risks and are vulnerable to implementation failure and ethical dilemmas. Future mental health work that necessitates the integration of interactive AI teammates will require understanding shared team goals (e.g., increasing the psychotherapist's fidelity to a given treatment protocol) and the ethical division of responsibilities between AI and humans (e.g., AI gathers data while the human makes a data-informed decision).

Central to understanding the ethical use of AI teammates is identifying how to sufficiently teach the psychotherapists about how the computational system works and its costs and potential benefits and risks. The *usability* of a worker-AI teaming workflow (e.g., see Fig. 1) will be evidenced by system features that the users (e.g., psychotherapist and supervisor) can learn easily. The system's *utility* will be demonstrated by showing that the users (e.g., treatment provider, trainer, administrators) have an accurate mental model of how the AI teammate judges the psychotherapist's treatment fidelity (Schelble et al., 2022; Wiese & Burke, 2019). Users must understand that the AI is trained on data from other patients and other psychotherapists and is then applied to the psychotherapist's performance. If users do not understand how AI systems work, there is an ethical dilemma of inappropriate trust in the system that diminishes the user's sense of responsibility. Many system design features can facilitate or dampen trust (Schaefer et al., 2016), such as the extent to which AI judgments can be used against the psychotherapist (Hirsch et al., 2018) and whether the judgment is contestable (Hirsch et al., 2017). The issue of contestability is central to moving AI from a *coach* (unidirectional feedback) to a *teammate* (bidirectional interaction). Ultimately, the ethical use of AI in mental health work may be enhanced by placing responsibility back on the psychotherapist to correct errors that might harm both the patient and the psychotherapist (Fiske et al., 2019).

Research Implications

Computational models have shown the potential for delivering unidirectionally, automatized evaluation of some aspects of therapy (Chen et al., 2021; Cummins et al., 2019; Ewbank et al., 2021; Flemotomos et al., 2021, 2022; Sharma et al., 2021; Tanana et al., 2021). Broadly speaking, research is needed that goes beyond an understanding of how psychotherapists *interact* with AI and address how they will *team* with AI. This is a critical distinction as the former represents the technology as a *tool* to get work done, while the latter emphasizes the technology as a *collaborative teammate* (O'Neill et al., 2022). Integrating a new synthetic teammate into mental health work will change how work gets done in prodigious and novel ways. First, teaming with AI will create new collaborative dynamics. Research consistently shows that integrating a new human teammate will change how the work gets done, the roles and responsibilities of each individual teammate, and how well the team performs together (Beus et al., 2014; Chen & Klimoski, 2003; Mathieu et al., 2017). These changes will occur when integrating an AI teammate into mental health work and create *new* collaborative dynamics and opportunities (Demir et al., 2018). Hence, it is critical to understand how the nature of the work itself will change after integrating AI teammates. The success of integrating this technology is contingent on both (1) the efficacy of the AI teammate and (2) overcoming the challenges and risks of implementation into existing systems.

Working in teams is an essential part of the modern workplace and mental health work is no exception (Robiner, 2006; Substance Abuse and Mental Health Services Administration, 2020). Introducing an AI teammate to the existing organizational system will change how current collaborative work is done (Demir et al., 2021; McNeese et al., 2018) by presenting new collaborative opportunities that are not currently part of mental health work. Consider the integration of AI teammates into the team-based tasks of upskilling psychotherapy skills. After completing their formal education, psychotherapists receive limited feedback throughout their careers (Amerikaner & Rose, 2012; Frank et al., 2020; Zukerman et al., 2023). Integrating an AI teammate would provide new opportunities for psychotherapists and AI teammates to learn together by reviewing session information and developing shared mental models (Schelble et al., 2022; Wiese & Burke, 2019).

Just like their human counterparts, an AI teammate must understand the existing collaborative dynamics to be an effective team member. Research should model current and future collaborative dynamics through a combination of inductive and deductive approaches. Observational studies can identify where current team dynamics could be augmented and where new worker-AI teaming could occur. Next, using established theoretical models of technology acceptance and adoption (e.g., Dingle et al., 2024; Hassan et al., 2024), researchers will need to forecast multilevel networks of collaborative dynamics that represent social, knowledge, resource, and task ties between collaborative actors and the strength of those ties. These approaches will not only identify the functional utility of the AI teammate (e.g., perceived usefulness, ease of use) but also the barriers to potential adoptions. It also opens the door to new research questions. For instance, psychotherapists tend to change their behaviors when being observed by other people (Ravid et al., 2020; Stanton & Weiss, 2000), but does this change when the observer is an AI teammate?

Innovations Needed for AI Teaming

Technology Design

Little is known about what psychotherapists expect or want in their future technology. Human–computer interactionists can use design inquiry (Zimmerman & Forlizzi, 2014) and design featuring (Kozubaev et al., 2020) methodologies to encourage users to consider technology that does not yet exist (Benjamin et al., 2021). Interviews with psychotherapists can elicit their ethical considerations, values, and feedback on including AI teammates in their work activity and what would constitute effective integration of AI teammates. Designers will also need to consider the role that various organizational structures play in designing the interface. For example, a clinical training environment might welcome the integration of invasive educational tools but a non-training environment with only licensed psychotherapists might be a context in which automated performance-monitoring is threatening, distracting, or offensive.

Novel interfaces and features need to be designed so that worker-AI teams can effectively communicate, coordinate, and adapt. System dashboards will need to mediate the visualization and exchange of information (i.e., communication) at the desired level of granularity. User-centered studies can identify what are the most effective ways to provide personalized feedback with actionable recommendations. Successful feedback requires that the psychotherapist can easily understand the rationale and justifications for given feedback (i.e., interpretability). Lastly, the form and delivery of feedback needs to be sensitive to user characteristics, such professional development level (i.e., novice versus expert). For example, an AI teammate that uses protocol-specific metrics can only help a user who knows the metrics and the theoretical frameworks that give meaning and importance to those metrics. A conceptualization of the various end-users can guide each step of design and development.

For end-users to cooperate with AI teammates, they may likely need to acquire knowledge regarding how AI teammates were built and how they work. Coordination in worker-AI teams will benefit from the end-user knowing enough about the system so that they can perceive it as reliable, directable, and intentional. For example, the psychotherapist might need to understand the basis of the AI teammate's feedback (e.g., a computational model trained on high-fidelity treatment session data). Correspondingly, the AI teammate might need to determine organizational hierarchies (e.g., supervisor versus trainee) and adapt the feedback weights appropriately, which may require the design of machine teaching interfaces (Simard et al., 2017).

Computational Innovations

In the specific domain of upskilling and sustaining psychotherapist skills, information science is presented with several challenging questions. Is it advantageous to develop a computation framework that assesses both the science of therapy (i.e., the content of protocols) and the art of therapy (i.e., the style in which these protocols are delivered)? How can the system provide actionable and specific feedback? How can systems incorporate a human-in-the-loop approach to update and rectify models? The vision of AI teammates to help learn psychotherapy hinges on the proliferation of computational innovations.

While significant progress is underway (Hirsch et al., 2017; Imel et al., 2019; Sharma et al., 2021), a pervasive challenge for information scientists will be collecting expertly annotated data based on real and effective therapy sessions to build computational models. One the most significant ethical risks is unintended bias in models resulting from biases in who is sampled, who is annotating, and who decides what data characteristics are important for the model. Some developers may use an auto-regressive controllable generative model to create a synthetic dataset for augmenting model training and evaluation (e.g., Chang et al., 2022). While recent work has established that this method can significantly improve training convergence (Rosenberg et al., 2019), its accuracy is largely unknown in the domain of psychotherapeutic encounters.

There are current advancements in assessing interaction styles and session dynamics using prosody features (e.g., turn-taking, back-channeling) and acoustic signals. Context-aware topic classifiers can detect specific protocol content (Gaut et al., 2015), active listening (Bodie et al., 2015), and rapport and therapeutic alliance (Goldberg et al., 2020; Yeh et al., 2019). Future research is needed on combining methods to assessment psychotherapy characteristics and identifying the incremental validity of each characteristic, which may or may not be protocol-specific (i.e., "common factors" such as empathy that are prescribed in most protocols).

The most striking upcoming innovations for worker-AI teaming will be bidirectional communication structures between the psychotherapists and the technology. This type of human-in-the-loop design allows the human teammates to provide feedback and potentially rectify model assessments and outcomes. Not only will this bidirectional learning system help engender trust in the AI teammate but it will also allow the worker-AI team to develop shared cognition, which is critical to accomplish their common goals (Schelble et al., 2022). Researchers will need to develop an interface to translate model outcomes into interpretable and actionable feedback. For example, if a model detects that a psychotherapist forgot to mention a treatment element, the interface can report an adherence issue and provide actionable recommendations. These interfaces can also convey behavioral signals captured throughout the session (e.g., using colors to indicate pitch intensity during the session). Through the use of an interface that enables editing and updating model outcomes, human teammates will be able to correct wrong assessments and incorrect feedback from models in a session, resulting in not only a more robust system but also providing a means to mitigate ethical concerns (e.g., transparency via explainability, justice via contestability).

Conclusions

Teaming with AI has potential transform the mental health workforce. One specific application of worker-AI teams is learning and sustaining psychotherapy skills (Imel et al., 2017). The ultimate impact of this emerging area of innovation is to efficiently upskill the workforce and, thereby, increase broad access to effective mental health treatment. Alongside potential benefits of teaming with AI, we highlight workplace, ethical, and research implications that need prioritization by practitioners, scientists, designers, and developers. Implications may differ between novice psychotherapists, who might use AI teammates to facilitate supervision, and established psychotherapists, who might use AI teammates for ongoing skill refinement. We provide broad recommendations for how to achieve technology innovations that are useful and usable. These developments will meet understandable levels skepticism and caution and all initial AI teammate integration should anticipate growing pains. The eventual success will require collaboration across all stakeholders and strict adherence to ethical principles.

Funding This work was supported by National Science Foundation Grant #2326144.

Data Availability No datasets were generated or analyzed during the current study.

Declarations

Ethical Approval Research with human or animal subjects was not conducted; therefore, ethical approval was not required.

Consent to Participate Research with human or animal subjects was not conducted; therefore, consent to participate was not required.

Consent for Publication Research with human or animal subjects was not conducted; therefore, consent to publish was not required.

Competing Interests The authors declare no competing interests.

References

Amerikaner, M., & Rose, T. (2012). Direct observation of psychology supervisees' clinical work: A snapshot of current practice. *The Clinical Supervisor, 31*, 61–80. https://doi.org/10.1080/07325223.2012.671721

Bash, H. L. & Stirman, S. W. (2020). Dissemination and implementation of evidence-based interventions for adults with PTSD (pp. 942–974). In J. G. Beck & D. M. Sloan (Eds.), *The Oxford handbook of traumatic stress disorders, second edition*. Oxford University Press.

Becker, E. M., Smith, A. M., & Jensen-Doss, A. (2013). Who's using treatment manuals? A national survey of practicing therapists. *Behaviour Research and Therapy, 51*(10), 706–710. https://doi.org/10.1016/j.brat.2013.07.008

Beidas, R. S., Edmunds, J. M., Marcus, S. C., & Kendall, P. C. (2012). Training and consultation to promote implementation of an empirically supported treatment: A randomized trial. *Professional Psychology: Research and Practice, 63*, 660–665. https://doi.org/10.1176/appi.ps.201100401

Békés, V., & Aafjes-van Doorn, K. (2020). Psychotherapists' attitudes toward online therapy during the COVID-19 pandemic. *Journal of Psychotherapy Integration, 30*(2), 238–247. https://doi.org/10.1037/int0000214

Benjamin, J. J., Berger, A., Merrill, N., & Pierce, J. (2021). Machine learning uncertainty as a design material: A post-phenomenological inquiry. *In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). https://doi.org/10.48550/arXiv.2101.04035

Berretta, S., Tausch, A., Ontrup, G., Gilles, B., Peifer, C., & Kluge, A. (2023). Defining human-AI teaming the human-centered way: A scoping review and network analysis. *Frontiers in Artificial Intelligence, 6*, 1250725. https://doi.org/10.3389/frai.2023.1250725

Beus, J. M., Jarrett, S. M., Taylor, A. B., & Wiese, C. W. (2014). Adjusting to new work teams: Testing work experience as a multidimensional resource for newcomers. *Journal of Organizational Behavior, 35*, 489–506. https://doi.org/10.1002/job.1903

Bodie, G. D., Vickery, A. J., Cannava, K., & Jones, S. M. (2015). The role of 'active listening' in informal helping conversations: Impact on perceptions of listener helpfulness, sensitivity, and supportiveness and discloser emotional improvement. *Western Journal of Communication, 79*(2), 151–173. https://doi.org/10.1080/10570314.2014.943429

Brennan, L. A., Brady, J. E., Drummond, K. L., Wiltsey-Stirman, S., Gutner, C. A., & Iverson, K. M. (2022). Mental health clinician perspectives regarding factors impacting implementation of evidence-based psychotherapies in Veterans Health Administration community-based outpatient clinics. *General Hospital Psychiatry, 75*, 54–60. https://doi.org/10.1016/j.genhosppsych.2022.01.001

Burton, M. S., Sherrill, A. M., Zwiebach, L. C., Fenlon, E. E., Rauch, S. A., & Rothbaum, B. O. (2023). Consultation competencies in prolonged exposure therapy for posttraumatic stress disorder. *Psychological Trauma: Theory, Research, Practice, and Policy, 15*(2), 279–286. https://doi.org/10.1037/tra0001115

Chang, J. H. R., Shrivastava, A., Koppula, H., Zhang, X., & Tuzel, O. (2022). Style equalization: Unsupervised learning of controllable generative sequence models. *In International Conference on Machine Learning* (pp. 2917–2937). https://doi.org/10.48550/arXiv.2110.02891

Chen, Z., Flemotomos, N., Ardulov, V., Creed, T. A., Imel, Z. E., Atkins, D. C., & Narayanan, S. (2021). Feature fusion strategies for end-to-end evaluation of cognitive behavior therapy sessions. *In 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society* (pp. 1836–1839). https://doi.org/10.1109/embc46164.2021.9629694

Chen, G., & Klimoski, R. J. (2003). The impact of expectations on newcomer performance in teams as mediated by work characteristics, social exchanges, and empowerment. *Academy of Management Journal, 46*(5), 591–607. https://doi.org/10.2307/30040651

Cho, E., Wood, P. K., Taylor, E. K., Hausman, E. M., Andrews, J. H., & Hawley, K. M. (2019). Evidence-based treatment strategies in youth mental health services: Results from a national survey of providers. *Administration and Policy in Mental Health and Mental Health Services Research, 46*, 71–81. https://doi.org/10.1007/s10488-018-0896-4

Cook, J. M., Biyanova, T., & Coyne, J. C. (2009a). Barriers to adoption of new treatments: An internet study of practicing community psychotherapists. *Administration and Policy in Mental Health and Mental Health Services Research, 36*, 83–90. https://doi.org/10.1007/s10488-008-0198-3

Cook, J. M., Schnurr, P. P., Biyanova, T., & Coyne, J. C. (2009b). Apples don't fall far from the tree: Influences on psychotherapists' adoption and sustained use of new therapies. *Psychiatric Services, 60*(5), 671–676. https://doi.org/10.1176/ps.60.5.671

Creed, T. A., Kuo, P. B., Oziel, R., Reich, D., Thomas, M., O'Connor, S., ... & Atkins, D. C. (2022). Knowledge and attitudes toward an artificial intelligence-based fidelity measurement in community cognitive behavioral therapy supervision. *Administration and Policy in Mental Health and Mental Health Services Research*, 1–14. https://doi.org/10.1007/s10488-021-01167-x

Cummins, R., Ewbank, M. P., Martin, A., Tablan, V., Catarino, A., & Blackwell, A. D. (2019). TIM: A tool for gaining insights into psychotherapy. *In The World Wide Web Conference* (pp. 3503–3506). https://doi.org/10.1145/3308558.3314128

Damschroder, L., Hall, C., Gillon, L., Reardon, C., Kelley, C., Sparks, J., & Lowery, J. (2015). The Consolidated Framework for Implementation Research (CFIR): Progress to date, tools and resources, and plans for the future. *Implementation Science, 10*, A12. https://doi.org/10.1186/1748-5908-10-S1-A12

Demir, M., Cooke, N. J., & Amazeen, P. G. (2018). A conceptual model of team dynamical behaviors and performance in human-autonomy teaming. *Cognitive Systems Research, 52*, 497–507. https://doi.org/10.1016/j.cogsys.2018.07.029

Demir, M., Likens, A. D., Cooke, N. J., Amazeen, P. G., & McNeese, N. J. (2019). Team coordination and effectiveness in human-autonomy teaming. *IEEE Transactions on Human-Machine Systems, 49*(2), 150–159. https://doi.org/10.1109/THMS.2018.2877482

Demir, M., McNeese, N. J., Gorman, J. C., Cooke, N. J., Myers, C. W., & Grimm, D. A. (2021). Exploration of teammate trust and interaction dynamics in human-autonomy teaming. *IEEE Transactions on Human-Machine Systems, 51*(6), 696–705. https://doi.org/10.1109/THMS.2021.3115058

Dingel, J., Kleine, A. K., Cecil, J., Sigl, A. L., Lermer, E., & Gaube, S. (2024). Predictors of health care practitioners' intention to use AI-enabled clinical decision support systems: Meta-analysis based on the unified theory of acceptance and use of technology. *Journal of Medical Internet Research, 26*, e57224. https://doi.org/10.2196/57224

Dores, A. R., Geraldo, A., Carvalho, I. P., & Barbosa, F. (2020). The use of new digital information and communication technologies in psychological counseling during the COVID-19 pandemic. *International Journal of Environmental Research and Public Health, 17*(20), 7663. https://doi.org/10.3390/ijerph17207663

Ewbank, M. P., Cummins, R., Tablan, V., Catarino, A., Buchholz, S., & Blackwell, A. D. (2021). Understanding the relationship between patient language and outcomes in internet-enabled cognitive behavioural therapy: A deep learning approach to automatic coding of session transcripts. *Psychotherapy Research, 31*(3), 300–312. https://doi.org/10.1080/10503307.2020.1788740

Fiske, A., Henningsen, P., & Buyx, A. (2019). Your robot therapist will see you now: Ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *Journal of Medical Internet Research, 21*(5), e13216. https://doi.org/10.2196/13216

Flemotomos, N., Martinez, V. R., Chen, Z., Creed, T. A., Atkins, D. C., & Narayanan, S. (2021). Automated quality assessment of cognitive behavioral therapy sessions through highly contextualized language representations. *PLoS ONE, 16*(10), e0258639. https://doi.org/10.1371/journal.pone.0258639

Flemotomos, N., Martinez, V. R., Chen, Z., Singla, K., Ardulov, V., Peri, R., ... & Narayanan, S. (2022). Automated evaluation of psychotherapy skills using speech and language technologies. *Behavior Research Methods, 54*(2), 690–711. https://doi.org/10.3758/s13428-021-01623-4

Frank, H. E., Becker-Haimes, E. M., & Kendall, P. C. (2020). Therapist training in evidence-based interventions for mental health: A systematic review of training approaches and outcomes. *Clinical Psychology: Science and Practice, 27*(3), 20. https://doi.org/10.1111/cpsp.12330

Gallo, K. P., & Barlow, D. H. (2012). Factors involved in psychotherapist adoption and non-adoption of evidence-based interventions in mental health. *Clinical Psychology: Science and Practice, 19*, 93–106. https://doi.org/10.1111/j.1468-2850.2012.01276.x

Gaut, G., Steyvers, M., Imel, Z. E., Atkins, D. C., & Smyth, P. (2015). Content coding of psychotherapy transcripts using labeled topic models. *IEEE Journal of Biomedical and Health Informatics, 21*(2), 476–487. https://doi.org/10.1109/jbhi.2015.2503985

Goldberg, S. B., Flemotomos, N., Martinez, V. R., Tanana, M. J., Kuo, P. B., Pace, B. T., ... & Atkins, D. C. (2020). Machine learning and natural language processing in psychotherapy research: Alliance as example use case. *Journal of Counseling Psychology, 67*(4), 438. https://doi.org/10.1037/cou0000382

Hassan, M., Kushniruk, A., & Borycki, E. (2024). Barriers to and facilitators of artificial intelligence adoption in health care: Scoping review. *JMIR Human Factors, 11*, e48633.

Health Resources and Services Administration. (2024). *Projected supply and demand for healthcare workers through 2036.* U.S. Department of Health and Human Services. https://data.hrsa.gov/topics/health-workforce/workforce-projections

Henrich, D., Glombiewski, J. A., & Scholten, S. (2023). Systematic review of training in cognitive-behavioral therapy: Summarizing effects, costs and techniques. *Clinical Psychology Review, 101*, 102266. https://doi.org/10.1016/j.cpr.2023.102266

Hidalgo, C. A., Orghian, D., Canals, J. A., De Almeida, F., & Martin, N. (2021). *How humans judge machines*. MIT Press.

Hirsch, T., Merced, K., Narayanan, S., Imel, Z. E., & Atkins, D. C. (2017). Designing contestability: Interaction design, machine learning, and mental health. *In Proceedings of the 2017 Conference on Designing Interactive Systems* (pp. 95–99). https://doi.org/10.1145/3064663.3064703

Hirsch, T., Soma, C., Merced, K., Kuo, P., Dembe, A., Caperton, D. D., ... & Imel, Z. E. (2018). "It's hard to argue with a computer": Investigating psychotherapists' attitudes towards automated evaluation. *In Proceedings of the 2018 Designing Interactive Systems Conference* (pp. 559–571). https://doi.org/10.1177/21677026231172694

Huang, L., Cooke, N. J., Gutzwiller, R. S., Berman, S., Chiou, E. K., Demir, M., & Zhang, W. (2021). Distributed dynamic team trust in human, artificial intelligence, and robot teaming. *In Trust in human-robot interaction* (pp. 301–319). Academic Press. https://doi.org/10.1016/B978-0-12-819472-0.00013-7

Imel, Z. E., Caperton, D. D., Tanana, M., & Atkins, D. C. (2017). Technology-enhanced human interaction in psychotherapy. *Journal of Counseling Psychology, 64*, 385–393. https://doi.org/10.1037/cou0000213

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

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Karlin, B. E., & Cross, G. (2014). From the laboratory to the therapy room: National dissemination and implementation of evidence-based psychotherapies in the U.S. Department of Veterans Affairs Health Care System. *American Psychologist, 69*, 19–33. https://doi.org/10.1037/a0033888

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 617–627. https://doi.org/10.1001/archpsyc.62.6.617

Koerner, K., Levy, J., & Dimeff, L. A. (2022). Using technology to train and sustain delivery of evidence-based practices. *Cognitive and Behavioral Practice, 29*, 41–49. https://doi.org/10.1016/j.cbpra.2020.12.007

Kozubaev, S., Elsden, C., Howell, N., Søndergaard, M. L. J., Merrill, N., Schulte, B., & Wong, R. Y. (2020). Expanding modes of reflection in design futuring. *In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–15). https://doi.org/10.1145/3313831.3376526

Lattie, E. G., Nicholas, J., Knapp, A. A., Skerl, J. J., Kaiser, S. M., & Mohr, D. C. (2020). Opportunities for and tensions surrounding the use of technology-enabled mental health services in community mental health care. *Administration and Policy in Mental Health and Mental Health Services Research, 47*, 138–149. 10.1007%2Fs10488–019–00979–2

Manríquez Roa, T., Biller-Andorno, N. & Trachsel, M. (2021). The ethics of artificial intelligence in psychotherapy. In M. Trachsel, J. Gaab, N. Biller-Andorno, S. Tekin, & J. Z. Sadler (Eds), *Oxford handbook of psychotherapy ethics*. Oxford University Press.

Mathieu, J. E., Hollenbeck, J. R., van Knippenberg, D., & Ilgen, D. R. (2017). A century of work teams in the Journal of Applied Psychology. *Journal of Applied Psychology, 102*, 452–467. https://doi.org/10.1037/apl0000128

McNeese, N. J., Demir, M., Cooke, N. J., & Myers, C. (2018). Teaming with a synthetic teammate: Insights into human-autonomy teaming. *Human Factors, 60*(2), 262–273. https://doi.org/10.1177/0018720817743223

McNeese, N. J., Demir, M., Chiou, E. K., & Cooke, N. J. (2021). Trust and team performance in human–autonomy teaming. *International Journal of Electronic Commerce, 25*(1), 51–72. https://doi.org/10.1080/10864415.2021.1846854

O'Neill, T., McNeese, N., Barron, A., & Schelble, B. (2022). Human–autonomy teaming: A review and analysis of the empirical literature. *Human Factors, 64*(5), 904–938. https://doi.org/10.1177/0018720820960865

Ravid, D. M., Tomczak, D. L., White, J. C. & Behrend, T. S. (2020). EPM 20/20: A review, framework, and research agenda for electronic performance monitoring. *Journal of Management, 46*, 100–126. https://psycnet.apa.org/doi/https://doi.org/10.1177/0149206319869435

Robiner, W. N. (2006). The mental health professions: Workforce supply and demand, issues, and challenges. *Clinical Psychology Review, 26*, 600–625. https://doi.org/10.1016/j.cpr.2006.05.002

Rosenberg, A., Zhang, Y., Ramabhadran, B., Jia, Y., Moreno, P., Wu, Y., & Wu, Z. (2019). Speech recognition with augmented synthesized speech. *In 2019 IEEE Automatic Speech Recognition and Understanding Workshop* (pp. 996–1002). https://doi.org/10.48550/arXiv.1909.11699

Rosero, A., Dinh, F., de Visser, E. J., Shaw, T. & Phillips, E. (2021). Two many cooks: Understanding dynamic human-agent team communication and perception using overcooked 2. *In Proceedings of AI-HRI Symposium*. https://doi.org/10.48550/arXiv.2110.03071

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In R. W. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 3–29). Ablex Publishing.

Sammons, M. T., VandenBos, G. R., & Martin, J. N. (2020). Psychological practice and the COVID-19 crisis: A rapid response survey. *Journal of Health Service Psychology, 46*(2), 51–57. https://doi.org/10.1007/s42843-020-00013-2

Schaefer, K. E., Chen, J. Y., Szalma, J. L., & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human Factors, 58*(3), 377–400. https://doi.org/10.1177/0018720816634228

Schelble, B. G., Flathmann, C., McNeese, N. J., Freeman, G., & Mallick, R. (2022). Let's think together!: Assessing shared mental models, performance, and trust in human-agent teams. *Proceedings of the ACM on Human-Computer Interaction, 6*, 1–29. https://doi.org/10.1145/3492832

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. In *Proceedings of the Web Conference 2021* (pp. 194–205). https://doi.org/10.48550/arXiv.2101.07714

Sherrill, A. M., Burton, M. S., Zwiebach, L., Fenlon, E., Rauch, S. A. M., & Rothbaum, B. O. (2021). Creating a national network of community-based consultants in prolonged exposure for PTSD: Outcomes and lessons learned from a consultant training program. *Psychological Trauma: Theory, Research, Practice, and Policy, 13*(8), 911–919. https://doi.org/10.1037/tra0001042

Sherrill, A. M., Wiese, C. W., Abdullah, S., & Arriaga, R. I. (2022). Overcoming clinician technophobia: What we learned from our mass exposure to telehealth during the COVID-19 pandemic. *Journal of Technology in Behavioral Science, 7*(4), 547–553. https://doi.org/10.1007/s41347-022-00273-3

Simard, P. Y., Amershi, S., Chickering, D. M., Pelton, A. E., Ghorashi, S., Meek, C., ... & Wernsing, J. (2017). Machine teaching: A new paradigm for building machine learning systems. *In Proceedings of arXiv*. https://doi.org/10.48550/arXiv.1707.06742

Speers, A. J., Bhullar, N., Cosh, S., & Wootton, B. M. (2022). Correlates of therapist drift in psychological practice: A systematic review of therapist characteristics. *Clinical Psychology Review, 93*, 102132. https://doi.org/10.1016/j.cpr.2022.102132

Stanton, J. M., & Weiss, E. M. (2000). Electronic monitoring in their own words: An exploratory study of employees' experiences with new types of surveillance. *Computers in Human Behavior, 16*, 423–440. https://doi.org/10.1016/S0747-5632(00)00018-2

Stowers, K., Brady, L. L., MacLellan, C., Wohleber, R., & Salas, E. (2021). Improving teamwork competencies in human-machine teams: Perspectives from team science. *Frontiers in Psychology, 12*, 590290. https://doi.org/10.3389/fpsyg.2021.590290

Substance Abuse and Mental Health Services Administration (SAMHSA). (2020). National Mental Health Services Survey (N-MHSS): On mental health treatment facilities.

Tanana, M. J., Soma, C. S., Kuo, P. B., Bertagnolli, N. M., Dembe, A., Pace, B. T., ... & Imel, Z. E. (2021). How do you feel?: Using natural language processing to automatically rate emotion in psychotherapy. *Behavior Research Methods*, 1–14. https://doi.org/10.3758/s13428-020-01531-z

Terjesen, M. D., & Del Vecchio, T. (Eds.). (2023). *Handbook of training and supervision in cognitive behavioral therapy*. Springer.

Valenstein-Mah, H., Greer, N., McKenzie, L., Hansen, L., Strom, T. Q., Wiltsey Stirman, S., ... & Kehle-Forbes, S. M. (2020). Effectiveness of training methods for delivery of evidence-based psychotherapies: A systematic review. *Implementation Science, 15*, 1–17. https://doi.org/10.1186/s13012-020-00998-w

Wiese, C. W., & Burke, C. S. (2019). Understanding team learning dynamics over time. *Frontiers in Psychology, 10*, 1417. https://doi.org/10.3389/fpsyg.2019.01417

Woods, D. D., Patterson, E. S., & Roth, E. M. (2002). Can we ever escape from data overload? A cognitive systems diagnosis. *Cognition Tech Work, 4*, 22–36. https://doi.org/10.1007/s101110200002

Xiao, B., Imel, Z. E., Georgiou, P. G., Atkins, D. C. & Narayanan, S. S. (2015). 'Rate my therapist': Automated detection of empathy in drug and alcohol counseling via speech and language processing. *PLoS ONE, 10*, e0143055. https://doi.org/10.1371/journal.pone.0143055

Yeh, S. L., Lin, Y. S., & Lee, C. C. (2019, May). An interaction-aware attention network for speech emotion recognition in spoken dialogs. *In ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 6685–6689). https://doi.org/10.1109/ICASSP.2019.8683293

Zhang, R., McNeese, N. J., Freeman, G., & Musick, G. (2021). "An ideal human" expectations of AI teammates in human-AI teaming. *Proceedings of the ACM on Human-Computer Interaction, 4*, 1–25. https://doi.org/10.1177/1071181322661007

Zimmerman, J., & Forlizzi, J. (2014). Research through design in HCI. In J. S. Olson & W. A. Kellogg (Eds.), *Ways of knowing in HCI* (pp. 167–189). New York, NY: Springer New York.

Zukerman, R. K., Schieffelin, A. J., & Del Vecchio, T. (2023). Supervisory processes in the context of dissemination and implementation science. *Handbook of training and supervision in cognitive behavioral therapy* (pp. 193–205). Springer International Publishing.