2020

Could a Robot Be Your Psychotherapist?

Benjamin Huston
University of Denver

Could a Robot Be Your Psychotherapist?

Abstract

As technology has advanced over the years, it has been integrated into psychotherapy and changed the way that people receive mental health care (Schopp, Demiris, & Glueckauf, 2006). Many of these advances, such as telehealth practices, were seen as unsustainable until the public Internet offered broader access to technology-based care in the 1990s (Schopp, Demiris, & Glueckauf, 2006). These technology-based practices have since grown in popularity and with a recent increase in telehealth practices, text-based therapies, and applications to aid in mental health practices, modern therapy looks very different than it did even ten years ago (Fiske, Henningsen, & Buyx, 2019).

One fairly new area of development is Artificial Intelligence (AI), herein defined as "computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages" (The Law Library of Congress, 2019). As AI advances, it is being implemented into more fields such as food service, transportation, personal assistants, and even healthcare (Fiske, Henningsen, & Buyx, 2019). As AI is further integrated into mental health, I believe it will become a common tool for therapists and eventually even operate as a replacement to human therapists. I do not believe that AI will ever be able to perform as well as a human, but I do believe such systems will be utilized regardless of whether the field of psychotherapy accepts these robotic "clinicians" as viable mental health providers. In this paper I will review literature related to the efficacy and ethicality of these existing technologies while also discussing potential solutions to the issues that the intersection of psychotherapy and AI will create in the near future.

Document Type

Doctoral Research Paper

Degree Name

Psy.D.

Department

Graduate School of Professional Psychology

First Advisor

Michael Karson

Second Advisor

Mark Aoyagi

Third Advisor

Helene Simons

Keywords

Artificial intelligence, A.I., Psychology, Pscyhotherapy, Telehealth, Teletherapy

Subject Categories

Artificial Intelligence and Robotics | Clinical Psychology | Other Psychology | Psychology

Could A Robot Be Your Psychotherapist?

A DOCTORAL PAPER

PRESENTED TO THE FACULTY OF THE

GRADUATE SCHOOL OF PROFESSIONAL PSYCHOLOGY

OFFICE OF GRADUATE STUDIES

UNIVERSITY OF DENVER

IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE

DOCTOR OF PSYCHOLOGY

BY

Benjamin Huston, MA

APPROVED: _____

Michael Karson, Ph.D., J.D., A.B.P.P., Chair

_____

Mark Aoyagi, Ph.D.

_____

Helene Simons, Psy.D.

Could a Robot Be Your Psychotherapist?

As technology has advanced over the years, it has been integrated into psychotherapy and changed the way that people receive mental health care (Schopp, Demiris, & Glueckauf, 2006). Many of these advances, such as telehealth practices, were seen as unsustainable until the public Internet offered broader access to technology-based care in the 1990s (Schopp, Demiris, & Glueckauf, 2006). These technology-based practices have since grown in popularity and with a recent increase in telehealth practices, text-based therapies, and applications to aid in mental health practices, modern therapy looks very different than it did even ten years ago (Fiske, Henningsen, & Buyx, 2019).

One fairly new area of development is Artificial Intelligence (AI), herein defined as "computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages" (The Law Library of Congress, 2019). As AI advances, it is being implemented into more fields such as food service, transportation, personal assistants, and even healthcare (Fiske, Henningsen, & Buyx, 2019). As AI is further integrated into mental health, I believe it will become a common tool for therapists and eventually even operate as a replacement to human therapists. I do not believe that AI will ever be able to perform as well as a human, but I do believe such systems will be utilized regardless of whether the field of psychotherapy accepts these robotic "clinicians" as viable mental health providers. In this paper I will review literature related to the efficacy and ethicality of these existing technologies while also discussing potential solutions to the issues that the intersection of psychotherapy and AI will create in the near future.

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

A Need for AI Therapists?

Many people do not utilize conventional mental health services due to barriers such as socioeconomic status, access to transportation, stigma, distance, and misinformation (Schopp, Demiris, & Glueckauf, 2006; Heckman et al., 2017). These barriers have demonstrated a need for telehealth developments in psychology to provide more accessibility as people are able to video chat with their therapists remotely. Despite a demonstrable need, the field of psychology, licensing boards, and insurance companies has been resistant to the implementation of telehealth practices (Schopp, Demiris, & Glueckauf, 2006).[1] This resistance comes from concerns related to quality of care, privacy and confidentiality, reimbursement, and patient safety. While some solutions and guidelines for these ethical concerns have been addressed by the American Psychological Association (APA, 2020), questions related to these issues are still unclear in many situations (Schopp, Demiris, & Glueckauf, 2006; Heckman et al., 2017). Fortunately, the field of teletherapy is regulated and to become credentialed in the field, therapists must undergo additional training (Center for Credentialing and Education, 2020). This means that, while problems still exist, professionals are working towards solutions that allow patients to receive the best quality of care possible when utilizing telehealth services. Additionally, while the nature of the telehealth-mediated therapeutic relationship is certainly different than conventional therapy, research shows that, for certain groups, telehealth interventions are generally effective (Varker et al., 2019).

This resistance to telehealth practices—combined with other barriers such as internet access, security, and technological proficiency—has limited the reach of teletherapy practices.

[1] It should be noted that the world was hit by the COVID-19 pandemic during the timeframe in which this paper was written. The resulting need for people to socially distance has greatly relaxed resistances to telepsychology. This sudden move towards an increase in teletherapy practice may continue following the pandemic. The virus has likely also created a greater demand for AI mental health support.

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

While teletherapy practices have still managed to grow in popularity, they remain inaccessible to large portions of the population. This has left a need for affordable, discrete, and easily accessible therapeutic interventions, fostering the development of AI solutions for mental health care. Major companies such as Amazon and Google have responded to this need and have had a hand in developing these solutions (Lock, 2018). With AI companions like Amazon's Alexa already existing in people's homes, a therapeutic tool programed into the AI would provide many people with easily accessed mental health support. This is exactly what Amazon and Google are currently working on with Mindscape, an application intended to help with everyday stressors through an AI companion that learns about its user, flags repeat problems, and offers advice and therapeutic interventions (Lock, 2018).

Mindscape is not alone in this development either. Other companies are developing their own AI "therapists" as companions to be used on either a computer or a smartphone, such as Woebot, Textpert, and MeetYourself (de Mello and de Souza, 2019; Fitzpatrick, Darcy & Vierhile, 2017; Textpert.ai. 2019). These applications are typically more affordable than conventional therapy, with some (such as Woebot) even being free to use. As it stands right now, these applications are in their infancy, and it is recommended that they are used as supplements or introductions to therapy (Fitzpatrick et al., 2017; Lock, 2018). As time goes on, and as developers continue to gather data from users, the hope is that these AI systems will evolve as "replacements" to conventional therapists.

Can an AI Provide Mental Health Care?

While many professionals believe that robots will never replace humans in roles that require social interaction, AI advancements have already shown strong potential such as taking on diagnostic roles in hospitals (Bottles, 2011). Such AI systems can ask questions about a

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

patient's symptoms, analyze these questions, and use medical literature to generate a differential diagnosis. While this approach to care may sound cold or aloof, Bottles (2011) noted that some patients actually prefer these robotic experts over human ones. In one case, a woman took her sick child to the emergency room and interacted with AI at a medical kiosk which briefly interviewed her about the child's symptoms. The system informed her that the child did not require immediate medical attention, and it made an appointment for the child at a later date. When interviewed about her experience, the woman remarked that the kiosk's system felt more compassionate about her child as compared with the triage nurses who were working in the ER (Bottles, 2011). While this is only one example, one can easily imagine there are already people (e.g., Millennials, Gen Z) who might be more comfortable utilizing and interacting with AI in professional as well as social settings.

Of course, using AI in a diagnostic role is one thing; using it for talk therapy is more complex and intricate. So, what would it take for an AI system to function autonomously as a therapist? Many AI experts would argue that the first step would be for the robot to pass the Turing test, a commonly used benchmark for measuring artificial intelligence (Warwick & Shah, 2015). Developer Alan Turing originally called it the "Imitation Game" (Turing, 1950). To pass this test a machine must be able to perform as a human during 30% or more of a five-minute conversation. While some consider this to be a hindrance to the way AI is understood and developed, it is considered by many experts to be an important way to gather evidence of machine thinking and practical testing for AIs (Warwick & Shah, 2015).

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

Some experts suggest that AI has already met the Turing benchmark. In a study by Warwick and Shah (2015), an AI system known as Eugene Goostman passed as a person during multiple five-minute text conversations 33% of the time, tested across 30 participants.[2] While the AI in this study was able to pass the Turing test, participants were still wary of the agents they were corresponding with, and they had difficulty communicating with them. These factors, along with the text-based nature of the conversations, would likely impact any attempt at building a therapeutic relationship on such a basis (Hull, 2015). However, it may be possible that awareness of the artificial nature of the AI would make it easier for users to connect with it. Fitzpatrick, et al. (2017) found that some individuals, such as those with social anxiety symptoms, might prefer to use AI because it removes the human variable.

Can You Bond with a Robot?

In a majority of the applications that exist today, the robotic nature of AI is made fully explicit. One of these AI programs, called Woebot, even reflects in its name the fact that it is a robot. Despite this the application has a sizable user base, reporting a couple hundred thousand users each month and two million messages sent a week (Rao, 2018). Woebot is designed to deliver brief cognitive behavioral therapy (CBT) interventions in the form of instant messenger conversations (Fitzpatrick et al., 2017). It currently exists as a free-to-use smartphone application that can be easily downloaded and set up, and thus sidesteps important barriers to mental health

[2] The participants communicated with one person and one AI system, for five minutes each, and were subsequently asked to identify if they had been addressing a person or an AI system. It should also be noted that a number of the human interactions failed the Turing test, which would raise the final score to 43% for interactions in which people were mistaken about whether they were dealing with humanity or machinery. It should also be noted that the AI sometimes identified itself as a young boy from the Ukraine as way to account for miscommunication, and this helped convince participants that they were communicating with a real person (Warwick & Shah, 2015).

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

services such as cost, transportation and distance. Additionally, because stigma is thought to be a primary reason that individuals refrain from seeking clinical services (Hunt & Eisenberg, 2010; Fitzpatrick et al., 2017). Woebot's creators have argued that AI therapy might be an attractive alternative, given its relative anonymity and rise in popularity.

While the empirical jury is out as whether Woebot and similar AI platforms are comparable to therapist-delivered CBT (e.g., Barak et al., 2008), Donkin et al. (2013) pointed out that these interventions are also characterized by a completion rate as low as 56%. The creators of Woebot theorized that part of the reason for the low completion rate for most internet interventions was the lack of a conversational element. In response, they attempted to design Woebot to mirror human interactions while integrating other evidenced-based recommendations for app development within CBT frameworks (Fitzpatrick et al., 2017). Recently, Fitzpatrick, et al. compared a test group utilizing Woebot to an informational control group. The study required participants to read an e-book which resembled many CBT-based sources in self-help. Fitzpatrick et al. hypothesized:

> A conversational agent built to incorporate both evidence-based guidelines for the development of mental health apps as well as hypothesized therapeutic process variables would be highly engaging, more acceptable, and would lead to greater reductions in symptoms of anxiety and depression relative to an information control group. (p. 8)

Participants in the control group had a significantly elevated attrition rate, which may show that Woebot's interactive design helped keep the interest of those in the test group. Results showed that participants in both groups experienced a significant reduction in anxiety symptoms, but only the test group showed a significant reduction in depression symptoms. Participants in the test group were much more engaged in the treatment and seemed to develop a relationship with

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

Woebot that many noted felt more empathic, even on the order of "a friend" (Fitzpatrick et al., 2017).

In order to describe what Woebot is like, I downloaded it myself to test it out. The application allows you to chat with Woebot on what resembles a texting screen. While the AI works out its reply, three little dots appear as if the robot were typing. It uses emoticons, slang, and texting language to communicate, and this does create the illusion that that you are texting with a real person. The application typically only allows multiple choice responses from the user, giving the conversation structure and direction but also making it more difficult to be as immersed in the conversation as you would with a real person. Woebot spent time introducing itself and explaining CBT before asking me how it could help. When it came time to discuss my problems, I decided to bring up some anxiety that I was experiencing. Woebot helped me to identify and reframe cognitive distortions and thinking errors. I was surprised to find that these techniques, introduced through my conversation with Woebot, actually helped to reduce my anxiety. Additionally, the AI would send me notifications about once daily. While I found this feature to be a bit annoying, I did occasionally respond to the notifications—showing that the application successfully pulls for adherence.

Anecdotal reports of Woebot have been offered by others. Brodwin (2018), viewed Woebot as an early but positive change to how people receive mental health treatment and noted how the AI provided some valuable insight, while feeling either like a conversation or a game. Brodwin pointed out that while Woebot was very clearly a machine, she praised it as an innovation that allows people to access mental health support in a new way. Torres (2017), on the other hand, was less impressed and felt that talking to a robot was "wasting time." Torres also tested Woebot's response to crisis by typing the word "suicide" into the chat box, to which

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

Woebot provided a crisis line number and quickly changed topics, with the suggestion to listen to some "relaxing music." As Torres asserted, and I agree, that response is entirely out of sync with what a human therapist would do.

At this point, Woebot's creators have asserted that it is meant to supplement rather than replace actual talk-therapy (Fitzpartick, et al. 2017), although it felt to me more like a rudimentary substitute. There is not any programming in place that allows Woebot to either pull from an individual's therapy sessions, or to prime them for what to discuss in therapy. It could operate as an alternative to a user's therapist when the therapist is unavailable, but it would not be able to assist the therapist directly for routine referrals or with emergency situations. Although it may perform better than traditional information-based apps, and there is some evidence that it is appealing to some individuals, the path to being on par with conventional therapy remains a ways off.

Programing the AI's "Mind"

To bring AI therapy up to par with conventional therapy, it would need to pass as human-like; that is, present with the "mind" of a human. Towards this goal, Shank et al. (2019) pulled from the research of Gray, Gray, and Wegner (2007) to describe the human mind in two dimensions: the agentic and experiential. *Agentic mind* relates to planning, outcome memories, communication, intention, and reasoning. *Experiential mind* relates to emotions, experiential memories, physical sensations, personality, and consciousness. According to Shank, et al., the agentic and experiential mind combine to form two separate categories of mind: the prototypical mind of human adults, and cryptominds, which are considered to have differentiated levels of intelligence. Cryptominds are perceived as lacking some degree in one of these dimensions, while a prototypical mind is seen to have both. Neurotypical adult humans have a prototypical

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

mind, while the cryptomind is reflective of babies, many animals, and AI. Babies, for example, have high experiential mind but low agentic mind, and AI has a moderate agentic mind and low experiential mind. An AI program with moderate agentic mind could easily teach therapeutic skills like those learned in CBT, but it would have difficulty applying those skills to its clients' lives (e.g., empathic listening and remembering important personal details) due to having a low experiential mind. Shank et al. asserted that any robot capable of passing as a human therapist would need to emulate and understand human emotions through both high experiential mind and high agentic mind.

Mori (1970) anticipated the hurdle of AI to achieving the combination of high agentic and experiential mind in describing *the effect of uncanny valley*—the difference between an entity presenting as perfectly human and an entity closely resembling human while being clearly non-human. Shank et al. (2019) have argued that long-term use of AI strengthens the perception of the AI's agentic and experiential mind, which blunts the uncanny valley and increases the user's satisfaction with the AI, particularly when a robot appears and sounds more human. While the creators of Woebot and similar text-based AI systems are seeking to do just that in order to form stronger connections with users, some (e.g., De Mello & de Souza, 2019) doubt this is possible over what they see as unreconcilable differences between AI and human intelligence.

More recent advances in AI development have focused on machine learning (ML), which requires extensive historical records so that it may make predictions and inferences (de Mello & de Souza, 2019). In this approach to AI, significant amounts of data are required to create a more human feeling AI system. The data needed to develop this type of AI come from a growing number of expanding sources, but will almost certainly need to come from users' personal history, which raises some important ethical issues around leaked or misused information. Users

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

of AI may agree to this risk; however, the cyberworld arguably presents greater risks to privacy (e.g., malware, spyware) than occurs in seeing a human therapist.

Is AI Therapy Ethical?

The companies currently developing AI alternatives to conventional therapy are largely unregulated (The Law Library of Congress, 2019). AI is a relatively new field, and many countries are rapidly developing AI technology in order to become leaders in the field, while its regulation is still in its infancy. Many countries, such as Canada and South Korea, have developed AI action plans that will eventually enact laws to regulate AI. However, it appears that no country has yet enacted these plans (The Law Library of Congress, 2019).

The lack of regulations for AI indicates a number of potential issues in ethical areas such as data protection and privacy, transparency, human oversight, surveillance, autonomous vehicles, and weapon systems. As it stands now, US regulations of AI are mainly concerned with autonomous—or self-driving—vehicles while AI in other areas including, psychotherapy, remains largely unregulated. As such, AI "therapists" and similar companions are not governed by the same rules (i.e., statutes and ethical codes) as mental health professionals. They do not need to be monitored by organizations like the American Psychological Association (APA), and they are not obligated to respond to ethical concerns like suicidal ideation or child abuse in the way that mandated reporters are (Fiske, Henningsen, & Buyx, 2019). The best they can typically do is provide the user with a crisis line to call when a crisis word triggers a response from their AI programming (Torres, 2017). This creates an ethical dilemma concerning AI therapists: they are unable to provide fully human responses to crises and would potentially prevent a user from seeking the help they need when they are at risk.

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

Additionally, these applications do not need to comply with the Health Insurance Portability and Accountability Act (Torres, 2017). Each application of programming for an AI therapist has its own privacy agreement, and some are not confidential. For example, using Woebot through Facebook requires consent to Facebook's privacy agreements, which includes allowing Facebook to read messages and utilize the data for their own means (see www.woebot.io/privacy, 2020). While this data mining may be used strictly for research and development of AI, it's access by company personnel also create opportunities for information leaks. Woebot's privacy policy acknowledges as much, accepting no liability for user data breeched by outside sources. This is not to say that Woebot takes no measures to protect user data, but the limits on this kind of privacy protection are difficult to find and understand—if they are read at all (it is commonly known that the vast majority of internet users do not)—compared to what a client would encounter in an actual therapist's mandatory disclosure statement and which the therapist can help explain if there are questions.

Further, the lack of regulation on AI also means that anyone, regardless of training or background in mental health, could develop an AI therapy application for public consumption. This presents the deep potential for scams and advertisement of unsafe or unproven clinical practices that professional associations (e.g., APA, American Counseling Association) and state departments of regulatory agencies (DORA) normally help to monitor.

Solutions to the AI "Therapist" Problem

Despite the inherent ethical issues wrought by AI therapy, there is little doubt that it will continue to grow in scale and popularity. Whether the effect on conventional mental health services might be positive (e.g., further normalize and encourage people to seek out conventional therapy), negative (e.g., people began to replace conventional therapy with an AI surrogate), or

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

neutral remains an empirical question, but it is clear the mental health industry needs to deal with these advances in AI therapy. I believe there are four ways that mental health professionals can meet this challenge: (a) integrating AI platforms into conventional therapy, (b) improving AI products through partnership with its developers, (c) improving regulation over AI therapy, and (d) expanding telepsychology practices. I will briefly discuss each of these below.

Integration of AI and Conventional Therapy

Given the state of technology and the associated ethical issues surrounding Woebot, and allied applications, I think they are at this time better thought of as adjuncts to conventional therapy. There are at least three ways that AI could serve as adjunctive tools. First, they might serve as a convenient, nonthreatening introduction to psychotherapy that helps to orient individuals to the types of responses and questions that an actual therapist might ask. Used in this manner, it would be important that the informed consent and feedback elements of AI are explicit about it not being a substitute to seeing a real therapist. de Mello and de Souza (2019) proposed something along these lines in investigating the use of AI to make psychotherapy briefer and more efficient. They utilized the platform MeetYourself (2018) to collect data and compile information over four sessions (using a programed series of questions that also provided patients with exercises to reframe perceptions and attitudes) that then generated reports, diagnoses, recommendations, genograms, and detailed social history data for the psychotherapist, who then took over for the next six sessions. According to de Mello and Souza, this system promotes a briefer therapeutic orientation that is likely to attract more patients to psychotherapists while also streamlining and simplifying the work for each therapist, possibly at a reduced cost to patients. Second, AI might be utilized as a kind of bibliotherapy, which is standard practice in many forms of psychotherapy and typically involves reading (e.g., books,

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

articles) as a therapeutic adjuvant. For example, instead of recommending a layperson's book or article on a particular form of treatment or clinical issue, a client could instead be directed to an appropriate AI platform for additional information or insight. Third, AI might be an excellent aftercare adjuvant, perhaps in place of "booster sessions" where important lessons and learnings are typically reinforced.

A potential downside to AI being utilized in these ways lies in therapists relying too heavily on them. For example, a therapist operating in the realm of private practice might be tempted to push the limits on their AI partner in order to expand their caseloads and make more money. This situation would likely lead to a lower quality of treatment and open up additional ethical issues around substandard practices of care (e.g., therapists "losing touch" with the content of their work, missing out on information or inferences they would have picked up on while doing the work themselves). An antidote to this concern might be a therapeutic companion application that could provide support during times that the therapist is unavailable while also relaying important information directly to the therapist related to homework, crises, setback, or any other material relevant to therapy. Applications like this already exist, such as Recovery Record (www.recoveryrecord.com/about, 2020), which is used to monitor goals, skills, meals, and heart rate for individuals who suffer from eating disorders. Such applications could be possibly integrated with an AI that could more directly respond to patient needs while keeping mental healthcare more tightly tethered to the human provider.

Improving Regulations for AI

Ethical concerns related to AI therapists are already concerning and will likely become more complex as AI technology advances. Areas such as liability, safety, evidence-based practice, and privacy will become ethically "grayer" as AI becomes better at imitating human

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

behaviors. To counter this complexity, AI, operating in the context of mental health care, will need clearer guidelines and regulations. I believe that applying the same liabilities faced by mental health care professionals to these AI could help to positively influence their development.

Safety, for example, would require that the AI have a more structured response to crisis words such as "suicide." This would be done by transferring the user to a crisis line and informing a living person that works for the company who then follows up with the individual. Privacy standards would create a more thorough disclosure statement for the users that outlines the ways in which client data is used and the ways in which it is vulnerable, while also ensuring that the company is doing everything within their power to protect their user's data. Competence and the use of evidence-based practice could be measured by an expert review of the application. These experts would determine whether or not the applications were accurately conveying specific evidenced based therapeutic approaches like CBT. The developers of the AI would be liable in the event of a crisis, and if a crisis is mishandled then the application would need to be pulled from the market and reviewed by a third party before being made available again. Additionally, the application needs to be clear about what services it is providing. If it detects that someone may need more specialized help (e.g., identifying someone as high risk for suicidal ideation) it should stop treatment and provide referrals to actual therapists. This would likely need to be done by having a trained professional, such as a crisis line operator or a therapist, contact the user by phone with the appropriate referrals in order to access for risk.

Partnership with AI Developers

Due to current regulations, any company or individual could create an AI for mental health care, resulting in a lack of oversight from psychotherapeutic practice and theory. While many companies developing AI "therapists" may be created by therapist or have therapist on

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

their teams, such as Woebot (www.woebot.io/about, 2020), information related to who they are consulting with and for what reasons is not readily available. This lack of oversight is concerning because it leaves room for these applications to be harmful to their users.

A solution to this problem would be to create a more official partnership between AI developers and psychotherapists. One way to achieve this would be for an organization such as APA to create an accreditation process, where they can train psychologists and psychotherapists to be consultants for AI. Accreditation would focus on similar issues as are required for licensure (e.g., liability, safety, evidenced based practice, and privacy), but extended to those issues that arise with online services and technology. The accrediting organization would create a set of standards required from these AI applications to receive accreditation, where the goals of these standards is to reduce the impact of these ethical concerns as much as possible.

Such standards would include criteria for responding to crisis situation, accurately conveying information, measuring user progress, informed consent, and limitations. They would require that a company has therapists on staff to contact and support users who are in crisis, while also providing local referrals as needed when they determine that a user needs more specialized help. The AI would also need to recognize and flag language that could indicate a user needs more support (e.g., suicidal or homicidal content). To determine that an AI is providing accurate information it should be reviewed by experts in specific areas. The company would also need to keep records of patient progress by asking users to complete evidenced based measures, such as the 9-Item Patient Health Questionnaire (PHQ-9), periodically throughout treatment. Consultants would then review the scores annually to identify any concerns related to the AI's quality of care. Additionally, these standards would require that a company provide a

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

clear informed consent agreement, providing information about liability, limitations, and privacy, before a user begins messaging the AI.

Consultants would help the AI developers to create their AI in a way that meets these standards. Following this, the accrediting organization would then review these AI mental health care applications and grant them a certification if they meet the set standards. While this accreditation would not be mandated, it would provide users with reassurance and likely lead to accredited applications being used in higher frequency than those that lack accreditation, providing an overall higher quality of care.

Embracing Teletherapy

Many of the factors that contribute to the demand and development of AI therapists could be met using other telehealth options. Teletherapy already allows patients to meet with therapists from remote locations, solving transportation problems and allowing patients to feel that there is a higher level of confidentiality (e.g., patients do not have to physically enter a mental health clinic). Even with these benefits, there are still concerns with affordability, insurance, and scheduling.

Some companies, like Talkspace and BetterHelp, have worked to make telehealth therapy solutions even more accessible. Talkspace, for example, offers on-demand therapy—either through video chat or through texting (Talkspace Online Therapy, 2020). This allows people to schedule therapy as they become available, and to seek help during times of crisis that may not warrant emergency services. Talkspace's text-based sessions may be cheaper and more accessible than other alternatives, but they have not been supported by insurance carriers (Hull, 2015). While some preliminary research on text-based psychotherapy has shown significant improvement in psychological well-being for 90% among participants, the same research has

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

also found a lower score for the therapeutic alliance available through text-based therapy when compared to traditional therapy (Hull, 2015). Thus, while such AI apps allow people to access therapy more flexibly, and possibly more anonymously at a reduced cost, the empirical jury remains in deliberation on how well these AI platforms can handle that even approximate what a real life therapist could. Nevertheless, to compete with artificial alternatives, and to overcome barriers preventing access to psychotherapy, psychotherapist will likely need to embrace these teletherapy options and find ways to improve them.

Among the ways for mental health providers to reduce the barriers that prevent patients from receiving teletherapy care is by promoting changes like that represented in the Medicare Telehealth Enhancement Act of 2005 (Schopp et al., 2006). This act was meant to eliminate restrictions on telehealth, allowing patients to receive reimbursement for care, and encouraging interstate licensure. It was designed to improve patient care and to save the state of Missouri money, by reducing the need for Medicaid travel reimbursement, which costs the sate $30 million a year. Promoting legislation like the Medicare Telehealth Enhancement Act would expand the reach of telehealth practices and change perceptions about teletherapy.

In order to provide the best quality of care for the greatest number of people, mental health providers must become involved in political processes that make teletherapy, and psychotherapy in general, more accessible. Legislature that encourages insurance coverage, covers transportation costs, and works to reduce barriers like interstate licensure would allow more people to access mental health services. In the removal of these barriers, a majority of people would choose to engage in services with real therapists, providing a higher quality of care for the patients, and a higher demand for therapists' services.

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

Summary and Conclusions

AI is a new field, and has not yet been thoroughly researched. Drawing on the most recent research I found that many studies relied heavily on survey and self-report measures from participants. As such, the extant empirical bucket has severe holes that will need to be patched and filled in order to better understand AI in the context of offering psychotherapy. Clearly, much research needs to be done to better understand the difference in quality of care between AI therapists and human therapists. Given the rapidity of technological development, there is little doubt that this will present an ever-evolving challenge to the field of mental health. As an individual highly interested and invested in both technology and human suffering, I hope I have endeavored above to highlight these challenges and suggest ways to address them.

References

American Psychological Association. (2020). Retrieved 5 March 2020, from https://www.apa.org/practice/guidelines/telepsychology

Barak, A., Hen, L., Boniel-Nissim, M., & Shapira, N. (2008). A comprehensive review and a meta-analysis of the effectiveness of Internet-based psychotherapeutic interventions. *Journal of Technology in Human Services, 26*(2–4), 109–160.

Bottles, K. (2011). Will patients trust sociable humanoid robots? Retrieved from https://www.kevinmd.com/blog/2011/08/patients-trust-sociable-humanoid-robots.html

Brodwin, E. (2018). I spent 2 weeks texting a bot about my anxiety—And found it to be surprisingly helpful. Retrieved 22 January 2020, from https://www.businessinsider.com/therapy-chatbot-depression-app-what-its-like-woebot-2018-1/

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

Brodwin, E. (2020). We're on the cusp of an explosive change in how we treat one of America's most ignored health problems. Retrieved 22 January 2020, from https://www.businessinsider.com.au/how-text-based-therapy-is-changing-how-we-treat-mental-health-2016-5

de Mello, F., & de Souza, S. (2019). Psychotherapy and artificial intelligence: A proposal for alignment. *Frontiers in Psychology, 10*.

Center for Credentialing and Education. (2020). Required Training | BC-TMH. Retrieved 5 March 2020, from https://www.cce-global.org/credentialing/bctmh/trainingc

Donkin, L., Hickie, I., Christensen, H., Naismith, S., Neal, B., Cockayne, N., & Glozier, N. (2013). Rethinking the dose-response relationship between usage and outcome in an online intervention for depression: Randomized controlled trial. *Journal of Medical Internet Research, 15*(10), 231.

Farley, F. (2013). Review of innovation, technology, psychology: A review of technology innovations for behavioral education. *Psychology of Aesthetics, Creativity, and the Arts, 7*(2), 210–211. https://doi.org/10.1037/a0030584

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. doi: 10.2196/13216

Fitzpatrick, K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health, 4*(2), 19.

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. Science,

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

315(5812), 619.

Harrison, A. M., & Goozee, R. (2014). Psych-related iPhone apps. *Journal of Mental Health, 23*(1), 48–50. doi:10.3109/09638237.2013.869575

Heckman, T. G., Heckman, B. D., Anderson, T., Lovejoy, T. I., Markowitz, J. C., Shen, Y., & Sutton, M. (2017). Tele-interpersonal psychotherapy acutely reduces depressive symptoms in depressed HIV-infected rural persons: A randomized clinical trial. *Behavioral Medicine*, *43*(4), 285–295. https://doi-org.du.idm.oclc.org/10.1080/08964289.2016.1160025

Hull, T. (2015). A preliminary study of talkspace's text-based psychotherapy. Retrieved 24 January 2020, from https://www.talkspace.com/online-therapy/wp-content/uploads/2015/06/Talkspace-Preliminary-Research-Study.pdf

Hunt, J., & Eisenberg, D. (2010). Mental health problems and help-seeking behavior among college students. *Journal of Adolescent Health, 46*(1), 3–10.

Johnston, B., Wheeler, L., Deuser, J., & Sousa, K. H. (2000). Outcomes of the Kaiser Permanente Tele-Home Health Research Project. *Archives of Family Medicine, 9*, 40–45.

Lock, H. (2018). This AI therapy skill for Alexa and Google aims to help with everyday mental health. [online] *The Ambient*. Available at: https://www.the-ambient.com/features/mindscape-alexa-therapy-skill-mental-health-1034 [Accessed 20 Jan. 2020].

Moon, K. M. (2014). Texting between client and therapist: An ethical dilemma or an effective tool? (Doctoral dissertation). Retrieved from https://search-ebscohost-

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

com.du.idm.oclc.org/login.aspx?direct=true&db=psyh&AN=2014-99100-410&site=ehost-live&scope=site

Mori, M. (1970). The uncanny valley. *Energy, 7*(4), 33–35.

Rao, A. (2018). *Woebot— Your AI Cognitive Behavioral Therapist: An Interview With Alison Darcy*. [online] Medium. Available at: <https://chatbotsmagazine.com/woebot-your-ai-cognitive-behavioral-therapist-an-interview-with-alison-darcy-b69ac238af45> [Accessed 25 May 2020].

Russel, S., & Norvig, P. (2009). *Artificial intelligence: A modern approach* (Third Edition). Saddle River, NJ: Prentice Hall.

Schopp, L., Demiris, G., & Glueckauf, R. (2006). Rural backwaters or front-runners? Rural telehealth in the vanguard of psychology practice. *Professional Psychology: Research and Practice, 37*(2), 165–173.

Schopp, L., Johnstone, B., & Merrell, D. (2000). Telehealth and neuropsychological assessment: New opportunities for psychologists. *Professional Psychology: Research and Practice, 31*, 179–183.

Shank, D., Graves, C., Gott, A., Gamez, P., & Rodriguez, S. (2019). Feeling our way to machine minds: People's emotions when perceiving mind in artificial intelligence. *Computers in Human Behavior, 98*, 256–266.

Talkspace Online Therapy (2020). Retrieved 12 February 2020, from https://www.talkspace.com/

Textpert.ai (2019). [online] Available at: https://www.textpert.ai/ [Accessed 20 Jan. 2020].

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.

The Law Library of Congress (2019). *Regulation of artificial intelligence in selected jurisdicitons*. Retrieved from https://www.loc.gov/law/help/artificial-intelligence/index.php

Torres, M. (2017). This new robot gives you a therapy session at your desk. [online] *Ladders | Business News & Career Advice*. Retrieved from https://www.theladders.com/career-advice/new-robot-gives-therapy-session-desk-want

Turing, A. M. (1950). Computing, machinery, and intelligence. *Mind, 59*(236), 433–460.

Varker, T., Brand, R., Ward, J., Terhaag, S. and Phelps, A., 2019. Efficacy of synchronous telepsychology interventions for people with anxiety, depression, posttraumatic stress disorder, and adjustment disorder: A rapid evidence assessment. *Psychological Services*, 16(4), pp.621-635.

Warwick, Kevin, & Huma Shah (2015). Can machines think? A report on Turing test experiments at the Royal Society. *Journal of Experimental & Theoretical Artificial Intelligence, 28*(6), 989–1007. doi:10.1080/0952813x.2015.1055826

Www.apa.org. (2020). *Ethical principles of psychologists and code of conduct*. [online] Available at: https://www.apa.org/ethics/code/ [Accessed 10 Feb. 2020].

Www.recoveryrecord.com (2020). Retrieved 26 April 2020, from https://www.recoveryrecord.com/about

Www.woebot.io (2020). Retrieved 26 April 2020, from https://www.woebot.io/about

Www.woebot.io (2020). Retrieved 24 January 2020, from https://www.woebot.io/privacy

Acknowledgement: I would like to express my thanks to my original chair, Ragnar Storaasli, Ph.D., who provided edits and feedback that led to the completion of this paper.