

# BROOKINGS

## COMMENTARY

# The wellness industry's risky embrace of AI-driven mental health care

Alexandrine Royer

October 14, 2021

If you need to treat anxiety in the future, odds are the treatment won't just be therapy, but also an algorithm. Across the mental-health industry, companies are rapidly building solutions for monitoring and treating mental-health issues that rely on just a phone or a wearable device. To do so, companies are relying on "affective computing" to detect and interpret human emotions. It's a field that's forecast to become a [\\$37 billion](#) industry by 2026, and as the COVID-19 pandemic has increasingly forced life online, affective computing has emerged as an attractive tool for governments and corporations to address an ongoing mental health crisis.

Despite a rush to build applications using it, emotionally intelligent computing remains in its infancy and is being introduced in the realm of therapeutic services as a fix-all solution without scientific validation nor public consent. Scientists [still disagree](#) over the over the nature of emotions and how they are felt and expressed among various populations, yet this uncertainty has been mostly disregarded by a wellness industry eager to profit on the digitalization of health care. If left unregulated, AI-based mental-health solutions risk creating new disparities in the provision of care as those who cannot afford in-person therapy will be referred to bot-powered therapists of uncertain quality.

The field of affective computing, also more commonly referred to as [emotion AI](#), is a subfield of computer science originating in the 1990s. Rosalind Picard, widely credited as one of its pioneers, [defined](#) affective computing as "computing that relates to,

arises from, or deliberately influences emotions." It involves the creation of technology that is [said to](#) recognize, express, and adapt to human emotions. Affective computer scientists [rely on](#) sensors, voice and sentiment analysis programs, computer vision, and ML techniques to capture and analyze physical cues, written text, and/or physiological signals. These tools are then used to detect emotional changes.

Start-ups and corporations are [now working](#) to apply this field of computer science to build technology that can predict and model human emotions for clinical therapies. Facial expressions, speech, gait, heartbeats, and even eye blinks are becoming profitable sources of data. [Companion Mx](#), for example, is a phone application that analyses users' voices to detect signs of anxiety. San-Francisco-based [Sentio Solutions](#) is combining physiological signals and automated interventions to help consumers manage their stress and anxiety. A sensory wristband monitors your sweat, skin temperature and blood flow, and, through a connected app, asks users to select how they are feeling from a series of labels, such as "distressed" or "content." Additional examples include the [Muse](#) EEG-powered headband, which guides users toward mindful meditation by providing live feedback on brain activity, and the [Apollo Neuro](#) ankle band, which monitors users' heart rate variability to emit vibrations that provide stress relief.

While wearable technologies remain costly for the average consumer, therapy can now come in the form of a free 30-second download. App-based conversational agents, such as [Woebot](#), are using emotion artificial intelligence to replicate the principles of cognitive behavioral therapy, a common method to treat depression, and to deliver advice regarding sleep, worry, and stress. Sentiment analysis used in chatbots combines sophisticated natural language processing (NLP) and machine learning techniques to determine the emotion expressed by the user. [Ellie](#), a virtual avatar therapist developed by the University of Southern California, can pick up on nonverbal cues and guide the conversation accordingly, such as by displaying an affirmative nod or a well-placed "hmmm." Though Ellie is not currently available to the wider public, it provides a hint of the future of virtual therapists.

In order to operate, artificial intelligence systems require a simplification of psychological models and neurobiological theories on the functions of emotions. Emotion AI cannot capture the diversity of human emotional experience and is often embedded with the programmer's own cultural bias. Voice inflections or gestures vary

from one population to another, and affective computer systems are likely to struggle to capture a diversity of human emotional experience. As the researchers Ruth Aylett and Ana Paiva [write](#), affective computing demands that “qualitative relationships must be quantified, a definite selection made from competing alternatives, and internal structures must be mapped onto software entities.” When qualitative emotions are coded into digital systems, developers use models of emotions that rest on shaky parameters. Emotions are no hard science, and the metrics produced by such software are at best an educated guess. Yet few developers are transparent about the serious limitations of their systems.

Emotional expressions manifested through physical changes also have [overlapping parameters](#). Single biological measures such as heart rate and skin conductance are not infallible indicators of emotional changes. A spiked heart rate may be the result of excitement, fear, or simply drinking a cup of coffee. There is still no [consensus](#) within the scientific community about physiological signal combinations that are the most relevant to emotion changes, as emotional experiences are highly individualized. The effectiveness of affective computing systems is seriously impeded by their [limited reliability, lack of specificity, and restricted generalizability](#).

The questionable psychological science behind some of these technologies is at times reminiscent of pseudo-sciences, such as physiognomy, which were rife with eugenicist and racist beliefs. In [Affective Computing](#), the 1997 book credited with outlining the framework for affective computing, Picard observed that “emotional or not, computers are not purely objective.” This lack of objectivity has complicated efforts to build affective computing systems without racial bias. Research by the scholar [Lauren Rhue](#) revealed that two top emotion AI systems assigned professional black basketball players more negative emotional scores than their white counterparts. After accusations of racial bias, recruitment company [HireVue](#) stopped using facial expressions to deduce an applicant’s emotional states and employability. Given the obvious risks for discrimination, AI Now called in 2019 for a ban on the use of affect-detecting technologies in decisions that can [“impact people’s lives and access to information”](#).

The COVID-19 pandemic exacerbated the need to improve already limited access to mental-health services amid reports of staggering increases in mental illnesses. In [June 2020](#), the U.S. Census Bureau reported that adults were three times more likely

to screen positive for depressive and/or anxiety disorders compared to statistics collected in 2019. Similar findings were reported by [the Centers for Disease Control and Prevention](#) <sup>7</sup>, with 11% of respondents admitting to suicidal ideation in the 30 days prior to completing a survey in June 2020. Adverse mental health conditions disproportionately affected young adults, Hispanic persons, Black persons, essential workers, and people who were receiving treatment for pre-existing psychiatric conditions. During this mental-health crisis, [Mental Health America](#) <sup>8</sup> estimated that 60% of individuals suffering from a mental illness went untreated in 2020.

To address this crisis, government officials loosened regulatory oversight of digital therapeutic solutions. In what was described as a bid to serve patients and protect healthcare workers, the [FDA](#) <sup>9</sup> announced in April 2020 it would expedite approval processes for digital solutions that provide services to individuals suffering from depression, anxiety, obsessive-compulsive disorder, and insomnia. The change in regulation was said to provide flexibility for software developers [designing devices](#) <sup>10</sup> for psychiatric disorders and general wellness, without requiring developers to state the different AI-ML-based techniques that power their systems. Consumers would therefore be unable to know whether, for example, their insomnia app was using sentiment analysis to track and monitor their moods.

By failing to provide instructions regarding the collection and management of emotion and mental health-sensitive data, the announcement demonstrated the FDA's neglect of patient privacy and data security. Whereas [traditional medical devices](#) <sup>11</sup> require testing, validation and recertification after software changes that could impact safety, digital devices tend to receive a light touch by the FDA. As noted by [Bauer et al.](#) <sup>12</sup>, very few medical apps and wearables are subject to FDA review, as the majority are classified as "minimal risk" and outside of the agency's enforcement. For example, under current regulation, mental health apps that are designed to assist users in self-managing their symptoms, but do not explicitly diagnose, are seen as posing "minimal risk" to consumers.

The growth of affective computing therapeutics is occurring simultaneously with the digitization of public-health interventions and the collection of data in self-tracking devices. Over the course of the pandemic, governments, and private companies [pumped funding](#) <sup>13</sup> into the rapid development of remote sensors, phone apps, and AI for quarantine enforcement, contact tracing, and health-status screening. Through the

[popularization of self-tracking applications](#)—many of which are already integrated into our personal devices—we have become accustomed to passive monitoring in our data-fied lives. We are nudged by our devices to record sleep, exercise, and eat to maximize physical and mental wellbeing. Tracking our emotions is a natural next step in the digital evolution of our lives—[Fitbit, for example, has](#) now added stress management to its devices. Yet few of us know where this data goes or what is done with it.

Digital products that rely on emotion AI attempt to solve the affordability and availability crisis of mental-health care. The cost of conventional face-to-face therapy remains high, ranging between \$65 to \$250 an hour for those without insurance based on the therapist directory [GoodTherapy.org](#). According to the [National Alliance on Mental Illness](#), nearly half of the 60 million individuals living with mental health conditions in the United States do not have access to treatment. Unlike a therapist, tech platforms are indefatigable and available to users 24/7.

People are turning to digital solutions at increasing rates to address mental-health issues. First-time downloads of the top 10 mental wellness apps in the United States reached [4 million](#) in April 2020, a 29% increase since January. In 2020, the Organisation for the Review of Care and Health Apps [found a](#) 437% increase in searches for relaxation apps, 422% for OCD, and 2483% in mindfulness apps. Evidence of their popularity beyond the pandemic is also reflected in the [growing number of corporations](#) offering digital mental-health tools to their employees. Research by [McKinsey](#) concludes that such tools can be used by corporations to reduce productivity losses due to employee burn out.

Rather than addressing the lack of mental-health resources, digital solutions may be creating new disparities in the provision of services. Digital devices that are said to help with emotion regulation such as the MUSE headband and the Apollo Neuro band cost \$250 and \$349, respectively. Individuals are thus encouraged to seek self-treatment through cheaper guided meditation and/or conversational bot-based applications. Even among smart-phone based services, many are hidden [behind pay-walls and hefty subscription fees](#) to access full content.

Disparities in health-care outcomes may be exacerbated by persistent questions about whether digital mental healthcare can live up to its analog forerunner. Artificial



intelligence is [not sophisticated enough](#) to replicate spontaneous, natural conversations of talk therapy, and [cognitive behavioral therapy](#) involves the recollection of detailed personal information and engrained beliefs since childhood—data points that cannot be acquired through sensors. Psychology is part science and part trained [intuition](#). As Dr. Adam Miner, a clinical psychologist at Stanford, [argues](#), “an AI system may capture a person’s voice and movement, which is likely related to a diagnosis like major depressive disorder. But without more context and judgement, crucial information can be left out”.

Most importantly, these technologies can operate without clinician oversight or other forms of human support. For many psychologists, the essential ingredient in effective therapies is the [therapeutic alliance](#) between the practitioner and the patient, but devices are not required to abide by [clinical safety](#) protocols that record the occurrence of adverse events. [A survey](#) of 69 apps for depression published in BMC Medicine found that only 7% included more than three suicide prevention strategies. Six of the apps examined failed to provide accurate information on suicide hotlines. Apps supplying incorrect information were reportedly downloaded more than [2 million times](#) through Google Play and the App Store.

As these technologies are being developed, there are no policies in place that dictate who has the right to our “emotion” data and what constitutes breaches of privacy. Inferences made by emotion recognition systems can reveal sensitive health information that poses risks to consumers. Depression detection by workplace software monitoring or wearables may cost individuals their sources of employment or lead to higher insurance premiums. [BetterHelp and Talkspace](#), two counseling apps that connect users to licensed therapists, were found to disclose sensitive information with third parties about users’ mental health history, sexual orientation, and suicidal thoughts.

Emotion AI systems fuel the wellness economy, in which the treatment of mental-health and behavioral issues are becoming a profitable business venture, despite a large portion of developers having no prior certification in therapeutic or counseling services. According to an estimate by the [American Psychological Association](#), there are currently more than 20,000 mental-health apps available to mobile users. One [study](#) revealed that only 2.08% of psychosocial and wellness mobile apps are backed by published, peer-reviewed evidence of efficacy.

Digital wellness tools tend to have high drop-out rates, as only a [small segment \(%20Systematic%20Search%20and%20Panel-Based%20Usage%20Analysis%20\(jmir.org\)\)](#) of users regularly follow treatment on the apps. An [Arean et al. ↗](#) study on self-guided mobile apps for depression found that 74% of registered participants ceased using the apps. These high attrition rates have stalled investigations into their long-term effectiveness and the consequences of mental health self-treatment through digital tools. As with other AI-related issues, non-White populations, who are underserved in psychological care, continue to be [underrepresented ↗](#) in the data used to research, develop, and deploy these tools.

These findings do not negate the ability of affective computing to provide promising medical and other healthcare developments. Affective computing has led to advances such as detecting spikes in heart rate in patients suffering from [chronic pain ↗](#), facial analysis to detect [stroke ↗](#), and speech analysis to [detect Parkinson's ↗](#).

Yet in the United States there remains no widely coordinated effort to regulate and evaluate digital mental-health resources and products that rely on affective computing techniques. Digital products marketed as therapies are [being deployed ↗](#) without adequate consideration of patients' access to technical resources and monitoring of vulnerable users. Few products provide specific guidance on their safety and privacy policies and whether data collected is shared with third parties. By being labelled as "wellness products," companies are not subject to the [Health Insurance Portability and Accountability Act ↗](#). In response, non-profit initiatives, such as the [Psyberguide ↗](#), have sought to rate apps by the credibility of their scientific protocols and transparency in privacy policies. But these initiatives are severely limited—and not a stand-in for government.

Beyond the limited proven effectiveness of these digital services, we must take a step back and evaluate how such technology risks deepening divides in the provision of care to already underserved populations. There are significant disparities in the United States when it comes to technological access and digital literacy. This limits the potential for users to make informed health choices and to consent to the use of their sensitive data. As digital solutions are cheap, scalable, and cost-efficient, segments of the population may have to rely on a substandard tier of service to address their mental health issues. Such trends also risk placing the responsibility for mental-health care on users rather than healthcare providers.

Mental-health technologies that rely on affective computing are jumping ahead of the science. Even emotion AI researchers are denouncing [overblown claims](#) made by companies and unsupported by scientific consensus. We do not have the sophistication of technology nor the confidence of science to guarantee the effectiveness of such digital solutions in addressing the mental health crisis. And at the very least, governmental regulation should push companies to be transparent about that.

***Alexandrine Royer** is a doctoral candidate studying the digital economy at the University of Cambridge, and a student fellow at the Leverhulme Centre for the Future of Intelligence.*

---

**The Brookings Institution is committed to quality, independence, and impact.**

We are supported by a [diverse array of funders \(/about-us/annual-report/\)](#). In line with our [values and policies \(/about-us/research-independence-and-integrity-policies/\)](#), each Brookings publication represents the sole views of its author(s).

Copyright 2025 The Brookings Institution