Speaker 1: This program contains mature discussion and language. Viewer discretion is advised.

Cinnamon Bless, Ph.D: In particular, I'll be focusing on the phenomenon of AI use for emotional support, for companionship out in the wild. What I'm calling out in the wild and what I mean by that is not under the auspices or guidance of a clinician. I look forward to talking about that this morning. By almost all accounts, the US is experiencing a youth mental health crisis that we've been in for a number of years now. In 2023, just two years ago, 18% of adolescents aged 12-17 who had a major depressive episode, but only half received mental health care. Many of you are clinicians may be familiar with that, but I invite you to take a moment and let that statistic sink in. These are some additional facts and figures from the National Alliance on Mental Illness, suggesting high distress levels and low access to care. Forty percent of US high school students had persistent sadness or hopelessness. In 2023, this number is up at 65% for LGBTQ+ students. Twenty percent of US high school students had serious thoughts of suicide tragically. Suicide is the second leading cause of death among people ages 10-24 years in the United States. When we think about young adults, individuals aged 18-25, only about 50% of individuals in that age group with a mental illness received treatment in 2024, so about a year ago. Refocusing on teens and like all people, we all experience episodes of feeling distressed, and we all use, as do teens, many sources of support out in the wild, outside of clinical care. Things like self reflection, turning to family and trusted adults. For teens, in particular, they may look towards institutional supports through their school, relational help seeking from professionals. For teens who feel comfortable and can have access, can afford access, they may seek help from professionals in person or seek informational help from credible experts or from Dr. Google, as we heard about earlier today, they may look towards their friends and peers, community resources like sports teams and other communities or cultural and spiritual practices through their churches, for instance. What I want to talk today about, however, is the fact that we've been learning more and more that teens are turning to AI based conversational agents for emotional support, really. I wouldn't argue, certainly, there's not a lot of rigorous evidence in this space yet, but I wouldn't argue that it's necessarily replaced these other sources of support, but I would say that I've mentioned the gaps, and I think that AI is being leveraged to fill the gaps. You can see here Harvard Business Review made the case early in 2025 that therapy and companionship moved from being much lower on the use case list in 2024 to becoming the top use case for generative AI in 2025. You can see here all of these headlines are from 2025 from a number of different news sources, suggesting that this is certainly becoming a phenomenon that is receiving a lot of attention in our society for both adults and teens alike. But what I want to point out, too, is that we are starting to see evidence that there's high use among youth in particular. You can see some headlines up here that are particularly focused on youth. The future is emotional AI, and Gen Z offers an early glimpse. I myself have a Gen Zier, and what's I think notable too, is that I think one of the first research studies came out late last year in November this JAMA Network open paper, really looking at the use of generative AI for mental health advice among US adolescents and young adults. This was a nationally representative survey of over 1,000 youth, 12-21 years and found that 13% reported using generative AI for mental health advice, and higher rates were noted from individuals at the older end of that age range, 22%. Among users, 65% sought advice monthly or more often than monthly. Among users, over 90% found the advice somewhat or very helpful, which I think is pretty interesting. Then of course, continuing my review of headlines, tragically, a number of very difficult and concerning outcomes have been reported in the media, including suicides which have been linked to the use of generative AI for mental health purposes. In fact, these are in many ways, reflect lawsuits that are ongoing with major AI companies, most commonly OpenAI. Then finally, anybody in here look to Common Sense Media for advice about what to allow their kids to look at. I'm somebody who does that. I see a couple of hands out there. In fact, Common Sense Media published a report recently. You can look at the range of their findings, but I think the top line one that I emphasize here is that the source that many parents turn to found that mental health support is one of the most common and in their perspective, most dangerous ways that teens use AI. I would argue that whether you see this trend as a natural disaster: fire, flood, drought, or tornado, perhaps, or you see it as just a ripple in the

march forward of technology, I would argue that the ecosystem has already changed. One thing I have been thinking quite a lot about with a number of my colleagues is what steps are needed at this point to really guide this space and foster the healthy pro social evolution of this landscape that now includes AI. What I'm going to do next for the middle part of my talk is convey some ideas, that's really the best word I have for them, ideas for promoting youth mental health in the age of AI, and I'm going to link each one of the five ideas that I present to you to some work that I've been involved with in various parts of my work. Those ideas are as follows. I've listed them up here. Listen to and reflect on the lived experiences of youth. Two, develop research agendas that align with youth interests. Three, assess shifting preferences among youth and adults while seeking to understand fundamental values around mental health support. Four, monitor harms, and five prepare the workforces needed to support youth mental health in the age of AI. We'll start with the first one, listen to and reflect on the lived experiences of youth. What I want to talk a little bit about is some work that I know I've had a chance to speak with some of you about. Early last year, I was fortunate to receive some funding through our Vice Chancellor for research through what's called the Convene and Influence Program here at UCSD. This program is really focused on having projects that advance the university's strategic themes. In this case, it was the themes of health and technology through work across campus and across departments. We formed a cross campus working group with faculty from public health, cognitive science, neuroscience, psychiatry. You can see several of my colleagues up here, philosophy, data science to pursue a number of things over the course of the year. What I want to talk most about here this morning is actually a public symposium, much like the one we're here at today that we held late last year in December of 2025. We were so fortunate, and I just want to give a shout out to my incredible colleague, Dr. Desiree Shapiro, who's an adolescent psychiatrist, and this was really her brainchild. But we were able in this meeting to host a youth panel on lived experience, asking youth what is appealing and concerning about AI for emotional support. We had five youth on the panel I think between the ages of, I want to say 14-19. We asked what factors might make using AI for emotional support appealing. What do you think are factors that might make young people avoid it or features that would make AI feel safe and supportive? This is not a research study, more just conveying to you what we heard from this panel of lived experience youth. We heard two things in response to the first question. Overwhelmingly, the accessibility and availability of AI. You can see here, when you said all you have to do is open up your laptop and you have something that you can use to vent your feelings. The availability and accessibility. Second, the ability to express feelings without judgment. This idea that you can say whatever you want, and nobody's going to judge you. That was an overwhelming thing. Sorry, this is cut off. When we asked about concerns, I think some of the concerns I heard from them were quite interesting. One I'll just amplify here is a concern around substitution of human connection with AI interaction and risk of echo chambers and sycophantic responses, as well as the last one that you can see here, number 5, that AI may provide temporary relief but may undermine the development of critical interpersonal and emotional regulation skills. Really thinking about how might repeated or high use of AI in this manner really affect the developing trajectory of youth. You can see here, one youth reflecting on that, saying, a real person is just so much more valuable because they can look you in the face and call you out on what you're lying to them about. Youth also mentioned they want safe and approachable adults in addition to dialogue about AI safety. These were the three areas that they mentioned in terms of what they wanted to see from AI. Improve AI credibility and transparency. I'll let you take a look at the quotes. Second, encourage adults to validate emotions and foster safe relationships. Like if you don't want us to turn to AI, then adults, you better get your stuff together and be the resource that we need. Then, third, increase dialogue about AI safety and privacy into schools, and I'll talk a little bit more about that in a moment. I really wanted to emphasize that first point, and then I'll go a little bit faster through the remaining four points, but want to mention them nonetheless. Second, develop research agendas. In particular, this is focused on third party researchers that align with youth interests. This is a commentary that was actually developed by a postdoc in my lab, Dr. Melissa Karnaze and really tries to articulate this case for the need for third party research on human sought emotional support from

AI. This is currently in press at Nature Human Behavior. The argument is like research researchers vantage point complements that of companies that are innovating and competing for market share and profits. Very likely underestimating risks, as well as governing bodies seeking to protect public, possibly overestimating risks as well as shaped by politics. Emphasizing the role for third party researchers. This is really just scratching the surface of what's included in this commentary. In fact, I think there's over 30 next steps offered for researchers in this paper. But you can see here the six reasons that are outlined, and I'll emphasize number 2, which is conversational AI tools offer novel emotional experiences. I think that's likely particularly relevant for teens and one key next step is to assess whether human relationships are undermined. Really, as well, I think there are many articulated reasons that align with the expressed concerns of youth. The third idea, assess shifting preferences while seeking to understand fundamental values. I would say this is for youth as well as for all people. This is a project that we pursued in the Center for Empathy and Technology about a year ago, and there's been a lot of actually data that's come out on this in the interim, trying to get a sense of attitudes towards AI therapy. We affectionately called this study the confident AI study. We surveyed a convenience sample of US Internet users about 130 in the fall of 2024, assessed uses of AI in the past year. Then attitudes towards a hypothetical scenario of a conversational AI model that was licensed to provide online therapy. You can see here, these numbers, I think, are fairly well aligned with what we're seeing in news reports and other surveys that are coming out in the gray literature suggesting about 10-20% of people use AI for some type of social, emotional support, and this was within the past year. Then we asked, what would you want from a hypothetical AI chatbot license to provide online therapy? You can see some of the thematic areas that emerged. I think adults, which is what we surveyed here, but I would add youth to this want insight, advice, non judgment. That's a big one and to feel better. I think they're looking for resources that will accomplish that. Then finally, we gave people a hypothetical scenario that involved the experience of burnout of a person named Alex and asked whether they would choose a human or AI therapist, and you can see the percentages here. It's about 60/40, 40% choosing a mental health chatbot. Then we also asked about what are your concerns with AI therapist? I think one of the interesting things for me, seeing the blue box and the things that are listed over there is you could easily ask somebody, what are you concerned about with a human therapist and you'd probably get a pretty similar list. I would say here, preferences may shift, but findings suggest that adults and youth want to feel understood, safe, not judged. I would say this is really the priority, not necessarily whether the source of support is human or AI. I would say, as somebody who's trained as a clinical psychologist and really taught that the therapeutic relationship is the cornerstone of the mechanism of intervention, it's quite interesting to see these findings emerge. I want to talk a little bit about monitoring harms, as well. Here, I'm going to amplify the work of one of my doctoral students Deepa Manjanatha. Just convey that both looking to the literature to understand the effectiveness of chatbot based interventions is obviously critically important, but monitoring of harms is critical, as well. This was a systematic review of studies on AI chatbots in adolescent mental health. We had seen some reviews of the literature for interventions, but none that were really focused on adolescents and none that had systematically evaluated each one of the studies for whether or not they systematically evaluated the harms. This project, after screening over 1,200 articles, included 14 studies and included youth as well as college students. They were mostly app based interventions, looking at the effectiveness in treating depression, anxiety, mood, affect, stress, well being, and body image. What's interesting here is that not for every outcome that they looked at, but roughly half of the studies reported did report statistically significant improvements in at least one target mental health domain. But only two out of the 14 studies monitored and then discussed harms, and those that did included worsening of depressive symptoms, mental health related emergency department visits. These were some of the harms that were reported in these very few studies that looked at this. I think really underscores the need to look at this. Then finally, I'll make some comments about, and I think this could be a much longer discussion, certainly among this group of professionals, but the need to prepare as well as care for the workforces needed to support youth AI in the AI age. I'm going to talk about they're a little bit disconnected, but two different

things. One is the need to prepare future clinicians for patients who are using. We've heard about many of them this morning, ways AI is leveraged within the health system. But I think there's also a need to prepare clinicians for patients who are utilizing AI for mental health purposes outside of the healthcare system and how that interplays with clinicians who may be seeing patients. This is a new elective that my colleague and I are offering just this quarter to really hopefully help prepare future health professionals for these discussions about LLMs with their patients and their colleagues. We try to do a variety of things in the course, and it will, of course, undoubtedly be updated next year when we've had a chance to offer it this year. Then the other thing I want to mention in terms of workforce, very briefly, the role of our online safety workforce, the trust and safety professionals, the people who moderate content online. It's a little bit of a different project, but I think this is work again, from my doctoral student Deepa Manjanatha. We surveyed well over 600 individuals who are professionally working as content moderators and tried to get a sense of various associations between mental health outcomes and different elements of their work and different ways in which they might be supported at work, by their coworkers, by their bosses. Do they work in person, do they work online, things like that. I'm just showing you here a snippet of this work. But as you can see here, we assessed, depression, anxiety, burnout, post traumatic stress. These individuals are exposed to the worst elements of what is on the Internet. These outcomes are very real risks for them. I think the point I want to make here is that with the acceleration of AI, it becomes much easier to generate harmful content. It can be done much faster in much larger volumes. I think this digital mental health workforce, becomes significantly more at risk, which I think puts everybody at risk, including youth who are actively engaged in online spaces. I'm going to take the minutes that I have left and just make some comments about broader trends and some takeaways. First, I'll just say that I do think there are some things from history that can inform our thinking about the future of generative AI for youth mental health. One of them is that there's been a longstanding interest in automating mental health treatment. Some of you may recall the development of Eliza, one of the very first chat bots that was developed to essentially simulate a Rogerian psychotherapist, who for those of you who may not know, Carl Rogers was incredible thought leader in the field of clinical psychology and really the founder of the idea of giving unconditional positive regard to patients as a way to offer therapy. You can imagine how AI chatbots might be very well situated to take up that unconditional positive regard, a framework of psychology versus, say there maybe other areas of psychology that are not as well suited. But I think this long history of interest in automating treatment, this is another step down that road. Similarly, there's the direct to consumer movement in medicine, and I think I put up here direct to consumer genomics. This was an area that I was quite involved in a decade or so ago, but really was controversial in terms of the idea that people could have direct access to their genetic information without the involvement of a healthcare provider. One way of thinking about the phenomenon that I've been talking about here this morning is this idea of direct access to this form of mental health support. I think some of the things that we were talking about in the late 2000 around direct to consumer health technologies, we might be able to use some of those frameworks to think about this phenomenon as well. Then I think in addition many people more recently that I talked to about this phenomenon often talk about the analogy with social media and youth mental health, which has been a very difficult space to think through and also been very intractable from a policy standpoint. In fact, in many ways, I'm certainly not an expert on the policy space around AI, but I think in many ways, it seems to me from my vantage point that we've actually had a lot more policy activity in the generative AI space than we really saw go forward in the social media space. You can see here our surgeon general, I think it was in 2003, that finally a report came out noting that there should really be a public health consideration for social media after over a decade of concern about this. Then I'll also just say, and this ties back into, I think some of the comments that we heard from youth. I think the impact of local level actions are perhaps sometimes underappreciated. We heard from youth that having more discourse in their schools, thinking about how they might engage with mental health or emotional support through schools and I had to have that discussion around AI within the school setting, I think could be fruitful. I'm just showing you this is actually in my group, we

also do some work on social media. This is a paper that was published by with my graduate student Alli O'Daffer where we offered some recommendations to include schools a bit more thoughtfully and intentionally when thinking about steps for social media. I think these were our recommendations for social media, but I think they could easily be very straightforwardly repurposed for this AI space. Then, finally, I think I don't need to tell this room that there's a need to balance AI safety with our ethical obligation to increase access to treatment. We could all, I'm sure, think of many ethical considerations related to the use of AI in the ways that I've been describing. But we also, in my opinion, have an obligation to leverage what I think could be a very promising technology for addressing the gaps that we've discussed. I think being able to really hold both of these things at once and determine a way forward is our ethical obligation. I'll just close with some high level takeaways for you all about even more meta than the five ideas that I presented. I think this is a very difficult space, and it's very fast moving, and our youth are changing. My kids are changing right before my eyes, and I'm sure that's the same for all of you. I think work needs to go across sectors, disciplines, professions, stakeholders, and generations. This touches on the need to listen to and reflect on the lived experiences of youth. I also think there's a need, certainly from a research standpoint to be agile, consider networked hubs in this space that bring in all the different stakeholders that I just mentioned versus more traditional structures that are perhaps more linear. I think these hubs may be what's needed to figure out how to have impact and then to amplify that impact. Then also, as I mentioned, balance safety with the ethical obligation to increase access to mental healthcare. Then finally, I would just say to stay curious and strive to understand what will likely be a constantly changing ecosystem. That is one of the things I strive for in the work that I do. I'd like to thank all of these folks and thank you for your time and attention.