Pacific Asia Journal of the Association for Information Systems
ISSN 1943-7544
Research Paper doi: 10.17705/1pais.12201 Volume 12, Issue 2 (2020)

Intelligent Conversational Agents in Mental Healthcare Services: A Thematic Analysis of User Perceptions

Ashish Viswanath Prakash[1,*], **Saini Das**[2]

[1]Indian Institute of Technology Kharagpur, India, ashish.viswanath@iitkgp.ac.in
[2]Indian Institute of Technology Kharagpur, India, saini@vgsom.iitkgp.ac.in

Abstract

Background: *The emerging Artificial Intelligence (AI) based Conversational Agents (CA) capable of delivering evidence-based psychotherapy presents a unique opportunity to solve longstanding issues such as social stigma and demand-supply imbalance associated with traditional mental health care services. However, the emerging literature points to several socio-ethical challenges which may act as inhibitors to the adoption in the minds of the consumers. We also observe a paucity of research focusing on determinants of adoption and use of AI-based CAs in mental healthcare. In this setting, this study aims to understand the factors influencing the adoption and use of Intelligent CAs in mental healthcare by examining the perceptions of actual users.*

Method: *The study followed a qualitative approach based on netnography and used a rigorous iterative thematic analysis of publicly available user reviews of popular mental health chatbots to develop a comprehensive framework of factors influencing the user's decision to adopt mental healthcare CA.*

Results: *We developed a comprehensive thematic map comprising of four main themes, namely, perceived risk, perceived benefits, trust, and perceived anthropomorphism, along with its 12 constituent subthemes that provides a visualization of the factors that govern the user's adoption and use of mental healthcare CA.*

Conclusions: *Insights from our research could guide future research on mental healthcare CA use behavior. Additionally, it could also aid designers in framing better design decisions that meet consumer expectations. Our research could also guide healthcare policymakers and regulators in integrating this technology into formal healthcare delivery systems.*

Keywords: Artificial Intelligence, Thematic Analysis, Mental Health Chatbots, Technology Adoption, Privacy Calculus, Anthropomorphism.

Citation: Prakash, A. V., & Das, S. (2020). Intelligent Conversational Agents in Mental Healthcare Services: A Thematic Analysis of User Perceptions. *Pacific Asia Journal of the Association for Information Systems, 12*(2), 1-34. https://doi.org/10.17705/1pais.12201

Introduction

Mental health care is in crisis world over. World Health Organization (2017) reports that one in four people in the world is affected by mental health disorders at some point in their lives. Mental health illness continues to be the single largest source of health-related economic burden worldwide (Lozano et al., 2012; Vigo et al., 2016; Whiteford et al., 2015). Common disorders, namely depression, and anxiety, with an estimated 322 million and 264 million people affected worldwide contribute most to this burden (Ginn & Horder, 2012). Despite the increasing burden, there is an acute shortage of mental health workforce globally (9 per 100,000 population) and particularly in the South East Asia region (2.5 per 100,000 population) (WHO, 2018). While there are effective and known treatment methods for these conditions, less than half of the affected receive such treatments (WHO, 2017). The lack of resources, scarcity of trained health care providers, and social stigma associated with the disease are identified to be the major barriers to effective care (WHO, 2017). There is increasing public demand on the health care systems worldwide to address the barriers and provide a convenient, cost-effective, and evidence-based solution to the affected (Naslund et al., 2017).

In this milieu, digital technology was posited as an elixir to solve long-standing problems related to social stigma and demand-supply imbalance in mental health care delivery. It was predicted to provide flexible solutions that are more accessible, cost-effective, and potentially less stigmatizing than the traditional model of mental health therapy (Hollis et al., 2018). Poor access to mental health care, thus created a demand for scalable and non-consumable approaches (Kazdin & Rabbitt, 2013). Following this, there has been an explosion of interest in the development of mobile health tools that could supplement traditional mental health treatment (Bakker et al., 2016). However, despite demonstrated therapeutic effectiveness in the trials comparable to that of a therapist (Barak et al., 2008; Merry et al., 2012), the usage and continuance of these digital mental healthcare tools were found to be relatively poor (Donkin et al., 2013). The inability of technology to sufficiently engage the patients (Darcy, 2017) and the failure of clinical trial outcomes to translate into improved patient outcomes in real-world settings (Gilbody et al., 2015; Mohr et al., 2017) are few important reasons identified for poor adherence and use.

The renascence in AI technology injects new hopes in the revival of digital mental health interventions. AI-based Conversational agents (CA), are software programs that are capable of interacting with users through natural language through a text or voice-based interface (Sarikaya, 2017). This technology is evolving rapidly and is presently used in digital assistants such as Siri, Cortana, Alexa, etc. and in customer interfaces across online retail, and banking (Microsoft Corporation, 2018). This technology is currently applied in the creation of an innovative variant of digital mental health intervention (mental health CAs), which has the potential to effect lasting impact on psychotherapeutic care (Bendig et al., 2019). The automated CAs can deliver evidence-based therapies by mimicking human interactions in an engaging and nonjudgmental way, thus addressing the issues of poor adherence, insufficient clinician availability, and stigma associated with mental health care delivery (Fitzpatrick et al., 2017; Molteni, 2017).

Several consumer trends have given rise to an urgent need to deepen the current understanding of CA-based mental health therapy. Firstly, a large number of people are using virtual mental healthcare services like ‘7cups of Tea’, which allows text-based counseling with human counselors instead of in-person visits to a therapist (Miner et al., 2017). Secondly, people have started having intimate conversations with text-based social bots like Xiaoice and Ruhh (Shum et al., 2018). Thirdly, consumers are now increasingly becoming familiar and confident with AI-enabled personal assistants on their smartphones and smart devices. These trends are indicative of increasing readiness of the consumers in using chatbot based therapy. The scale of user downloads and reviews of mental health chatbots further reveals the emerging interest in these self-help tools.

DOI: 10.17705/1pais.12201

Amidst, the promised benefits and growing user enthusiasm, the emergence of fully automated agents for mental health therapy, has given rise to debates on various social and ethical issues (Kretzschmar et al., 2019; Martinez-Martin & Kreitmair, 2018). Questions are raised about the privacy, clinical efficacy, patient safety and accountability (Bhugra et al., 2017; Kretzschmar et al., 2019; Martinez-Martin & Kreitmair, 2018; Molteni, 2017) of CA-based mental health therapy. We believe several of these issues will act as probable inhibitors to the consumer adoption of these applications.

Moreover, academic research on mental health CAs has been largely focused on technical development, design characteristics, assessment of efficacy, and ethical challenges (Kretzschmar et al., 2019; Martinez-Martin, & Kreitmair 2018). It is observed that considerations on patient perceptions of potential benefits and risks of mental health chatbots are largely overlooked (Bendig et al., 2019; Morris et al., 2018). Overall there is a paucity of research addressing the individual adoption and use of CA in the context of healthcare and particularly with respect to mental health CAs. There is hardly any research focusing on a comprehensive analysis of enablers and inhibitors of mental health CA adoption from a theoretical viewpoint.

Therefore, to overcome barriers and to improve the delivery of AI chatbot based digital mental healthcare services and thereby the patient outcomes, it is important to explain how and why mental healthcare patients would use a CA for mental health support. Hence, the purpose of this research is to explore user perceptions and develop a theoretical model that will provide a basis to examine the acceptance of automated CA for mental healthcare support. Accordingly, we propose the following research question:

RQ: What are the factors influencing consumer's adoption and use of automated conversational agents providing mental healthcare services?

From a methodology perspective, the research uses a qualitative approach based on netnography (Kozinets, 2002) and thematic analysis (Braun & Clarke, 2006) to explore the relevant factors from the publicly posted user reviews (Google Playstore) of two popular mental health CAs. Based on a thorough thematic analysis of the reviews, we propose a thematic map of determinants explaining the consumer (patient) adoption of mental health CAs.

The contributions of this study are threefold. First, it throws light on the user's perception of emerging AI-based CAs providing mental health therapy and identifies critical issues pertaining to adoption from the user's perspective. Second, the research proposes a comprehensive thematic map of factors governing the user acceptance of AI-based mental healthcare CAs that could be tested using quantitative methods in the future. Finally, the implications for the practice offered by the study could aid designers of mental health CAs in developing new services or streamlining their existing service offerings to meet customer expectations. It will also provide insights to healthcare policymakers and regulators, in integrating these systems into formal healthcare delivery systems.

The rest of the paper is structured as follows, the following section reviews the background and related literature. The research methodology adopted for this study is elaborated subsequently. Thereafter the results are presented. In the final section of the paper detailed discussion about the major findings, its implication for research and practice, and the limitations & the scope for future research are presented.

Literature

A thorough literature review was carried out to identify the knowledge gaps in the literature and problematize research opportunities in the emerging area of user acceptance of automated mental health therapy. The review is presented under five subsections.

AI-based Conversational Agents

Artificial Intelligence, often known as 'non-biological intelligence' (Tegmark, 2017), is an umbrella term that represents 'the science and engineering of making intelligent machines' (McCarthy, 2007). Applications of AI ranges from backend recommender algorithms in online retail, to automated medical diagnostics to fully autonomous cars (Hengstler et al., 2016). A highly valued development in the area is the automated CAs, which are capable of understanding and responding to users in natural languages (Sarikaya, 2017). This technology has rapidly evolved in its capabilities from its humble beginnings of using keyword matching techniques (Shawar & Atwell, 2007) to real-time voice-based user interfaces that use sophisticated natural language processing (NLP) algorithms, which are capable of handling multiple languages. Consumer applications of this technology have progressed from smartphone interfaces (Apple Siri, Microsoft Cortana, etc.) to stand-alone smart devices (Alexa and Google Home). CAs are now extensively deployed in customer interfaces to help users complete specific tasks in online retail and banking (Microsoft Corporation, 2018).

With respect to the academic developments in the chatbots, we observe that most of the current research on chatbots are focused on the technical aspects of the technology (Sheehan, 2018) or related to the human-like characteristics of the bot (Hill et al., 2015; Schuetzler et al., 2018). We observe that literature on the consumer acceptance and use of CA is still in nascence. There are only a handful of studies on user acceptance of chatbot, e.g., pertaining to intelligent personal assistants (Han & Yang, 2018), tourism (Melián-González, et al., 2019), branding (Zarouali et al., 2018), and banking (Payne et al., 2018). Furthermore, few other related studies discussed issues pertaining to information disclosure (Schroeder & Schroeder 2018), privacy concerns (Saffarizadeh et al., 2017), user trust (Elson et al., 2018; Saffarizadeh et al., 2017), customer experience (Trivedi, 2019), and customer satisfaction (Chung et al., 2018). We notice a paucity of user acceptance studies that address the individual adoption and use of CA in the context of healthcare. There is hardly any research exploring health chatbot acceptability and motivations for its use from the patient's perspective.

Applications of CA in Mental Healthcare

Conversational AI is changing mental healthcare delivery. Developers are trying to automate certain aspects of mental health assessments and treatment by using conversational AI (Miners et al., 2017). Primarily, CAs are scalable and non-consumable interventions, which makes them an attractive alternative to other forms of digital healthcare options, which are inherently constrained by the time and attention of the clinician (Kazdin & Rabbitt, 2013). They also provide greater interactivity, flexibility, and control over the content and duration of the therapy (Gaffney et al., 2019). Additionally, it is also known that users respond and connect to CAs in social ways, and thus eliciting honest disclosures (Lucas et al., 2014).

Promising areas for the use of CAs in mental health context could be for support for diagnosis, prevention, treatment, and follow up of psychological problems (Bird et al., 2018; D'Alfonso et al., 2017; Huang et al., 2015). Recent research shows that suicidal ideation or behavior can be effectively detected from social media postings using AI (De Choudhury et al., 2016). Then, for instance, chatbots may automatically remind users of nearby mental healthcare services for assistance. Concerning treatment, CAs could provide resources for the treatment of psychological problems that participants may deal with on their own. Chatbots may also aid in

https://aisel.aisnet.org/pajais/vol12/iss2/1
DOI: 10.17705/1pais.12201

complementing the psychotherapy interventions, stabilizing interventions effects, and thus reducing the probability of relapse (D'Alfonso et al., 2017).

Globally a few applications are providing a variety of psychotherapy services, namely Woebot (English), Wysa (English), Joy (English), Shim (Swedish), Karim (Arabic), Emma (Dutch), Sabori (Japanese), etc. (Bendig et al., 2019). These CAs offer value propositions such as relief from stress, anxiety, and depression or general psychological wellbeing or emotions, mood regulation, or support behavioral changes. They are generally based on widely-accepted evidence-based psychotherapeutic approaches such as CBT or DBT or newer methods like mindfulness (Bendig et al., 2019). The unique characteristic features of CAs used for mental healthcare are summarized in figure 1.

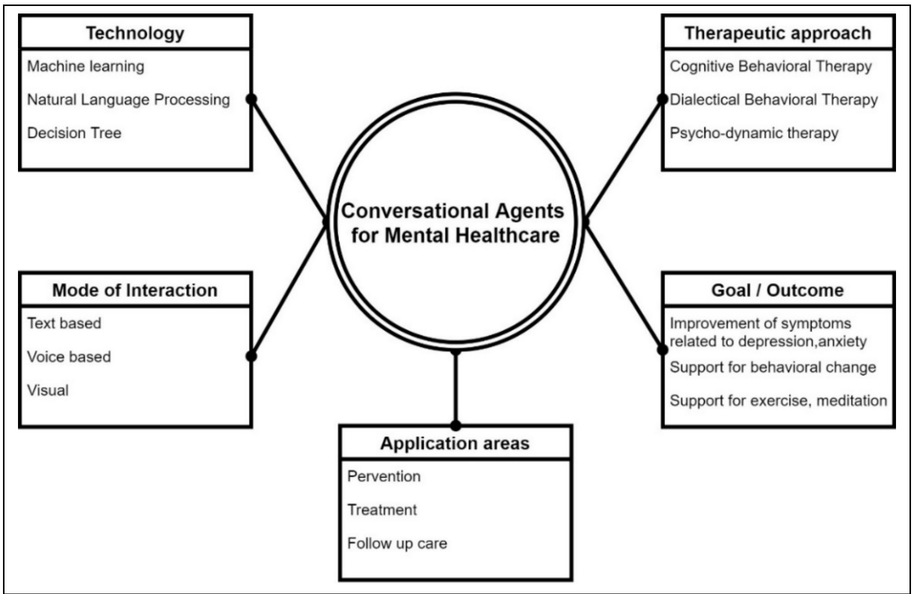


Figure 1 - Characteristics of CA for Mental Healthcare

Notwithstanding the promised benefits, certain issues indicate that the rapid diffusion of mental health CAs could be a considerable challenge. Firstly, currently, CAs are not sophisticated enough to recreate the richness of a conversation with a human therapist (Gaffney et al., 2019). Secondly, although the current clinical evidence on the efficacy of CA interventions appears promising, more evidence is required to demonstrate its equivalence to other treatment modalities to win the acceptance of users and clinicians (Gaffney et al., 2019). Thirdly research on ethical issues raises questions on privacy, efficacy, safety, and accountability of CA-based mental health therapy (Martinez-Martin & Kreitmair, 2018; Kretzschmar et al., 2019).

Research on the Adoption of Digital Mental Healthcare Services

We explored the prior research on the adoption of digital mental healthcare to understand the knowledge gaps and to identify the potential factors that could play a role in mental health CA adoption decisions. Our analysis of the literature reveals that academic interest in mental health CA is predominantly centered around the technical development of CAs, design characteristics, assessment of efficacy (Bendig et al., 2019; Gaffney et al., 2019) and ethical challenges (Kretzschmar et al., 2019; Martinez-Martin, & Kreitmair). It is observed that considerations on patient perceptions of potential benefits and risks of mental health chatbots are largely overlooked themes (Bendig et al., 2019).

Table 1 summarizes the prior research on the adoption and use of digital mental healthcare services. We observe that studies related to adoption and use of digital mental healthcare itself is scant, and the academic interest in the specific field of mental health CA usage is emerging. We observe that studies on adoption and use of digital mental healthcare services are predominantly from medical journals and were mostly exploratory and descriptive in nature without any sound theoretical foundation. Most of them followed a qualitative approach except for Huang and Bashir (2017), Torous et al., (2018) and Lipschitz et al., (2019).

Table 1 - Summary of Related Literature on Adoption and Use of Digital Mental Healthcare

Author (year)	Context	Findings related to factors affecting adoption/use	Methodology	Theoretical paradigm
Huang and Bashir (2017)	Anxiety apps	Information cues (App Price, Rating, & Review, App Permission and Category, App Ranking, Title)	Nonparametric regression technique	NIL
Fitzpatrik et al., (2017)	Mental health Chatbots	Identified best (check-in, personality, learning, conversation) and worst (process violation, technical problem, content) things about the experience of using mental health chatbot	Randomized Controlled Trial, Thematic analysis of Open-ended questions	NIL
Schueller et al., (2018)	Mental Health apps	Ease of use, aesthetics, and individual experience, features, trustworthiness of the source	Thematic analysis, Focus group discussion	NIL
Torous et al., (2018)	Mental Health apps	Privacy, Accuracy, set up, usability, Time-saving, cost-saving	Questionnaire-based study	NIL
Connolly et al., (2018)	Mobile app for mental healthcare	Treatment effectiveness, Ease of use, Culture and identity, connectivity barriers	Thematic analysis, Qualitative interviews	Technology acceptance theories in general
Fulmer et al., (2018)	Mental health Chatbots	Identified most and least favored features of chatbot therapy	Randomized Controlled Trial, Thematic analysis of Open-ended questions	NIL
Lipschitz et al., (2019)	Mobile Apps for Depression and Anxiety	Reasons for nonuse were lack of proof of efficacy, data privacy concerns, lack of personalization, and difficulty in use	Questionnaire-based study	NIL
Kretzschmar et al., (2019)	Mental health Chatbots	Identified ethical issues namely, privacy and confidentiality, safety, efficacy related to use of CA for mental healthcare	Focus group discussion	NIL
This research	*Mental health Chatbots*	*Factors influencing the adoption and use of AI-based CA for mental health support*	*Thematic analysis of user reviews*	*UTUAT2, Privacy Calculus, Trust Theory, Theory of Anthropomorphism*

From the Table 1, we observe that prior research on AI mental health chatbots (Fitzpatrik et al., 2017; Fulmer et al., 2018; Kretzschmar et al., 2019) lacks theoretical basis and were focused on identifying consumer's perspectives about ethical issues related to the use or favored features of chatbot therapy without any deliberations on how these factors would determine adoption and use. Of which the studies Fitzpatrik et al., (2017) and Fulmer et al.,

(2018) were primarily Randomized Controlled Trials on effectiveness but also additionally included a small section evaluating open-ended questions about user perceptions. Also, we observe that these studies did not use actual mental health patients (but used student/non-patient samples) to arrive at their conclusions. More importantly, very little research has been done on the user's perceptions of CAs for mental healthcare support (Morris et al., 2018), and there is hardly any research focusing on determinants of adoption of AI-based automated chatbots in mental healthcare.

After a thorough review, we observe the following research gaps in the literature: (1) There is a paucity of studies that address the individual adoption and use of CA in the context of healthcare. (2) There is lack of research on patient perceptions of potential benefits and risks of mental health chatbots (Bendig et al., 2019). (3) The studies related to adoption and use of digital mental healthcare, in general, lacks theoretical basis. (4) There is hardly any research focusing on a comprehensive analysis of enablers and inhibitors of mental health CA adoption using theoretical perspective. This study tries to address the above-mentioned research gaps by exploring the factors influencing consumer's decision to use automated CAs providing mental health care services. Furthermore, drawing upon the insights from prior research, we also presume that issues like privacy concerns, safety, ease of use, trust, and efficacy of the intervention may play an essential role in determining the adoption behavior of Mental health CAs.

Theories on Technology Acceptance

To address the identified knowledge gaps in the literature and thereby the issue of adoption challenge of Mental health CAs, we built on the technology acceptance research. Researchers from consumer marketing, behavioral psychology and IS has used several disjointed theoretical approaches to explain the acceptance behavior namely, Theory of reasoned action (TRA), Theory of planned behavior (TPB), Technology Acceptance Model (TAM), Innovation Diffusion Theory (IDT), Motivational Model (MM), Model of PC Utilization (MPCU), Combined TAM and TPB (C-TAM-TPB), Social Cognitive Theory (SCT), etc. (Venkatesh et al., 2003).

Among these, the Technology Acceptance Model (TAM), an adaptation of TRA, introduced by Davis (1989), is one of the most popular theories in IS that has been applied to explain individual technology adoption in a variety of settings (Venkatesh et al., 2003). TAM postulates that two specific perceptions, namely, perceived usefulness (PU) and perceived ease of use (PEOU), determine the intention to use a target technology and intentions, in turn, predict the usage behavior. As per the model, PU is defined as "the degree to which a person believes that using a particular system would enhance his or her job," while PEOU is defined as "the degree to which a person believes that using a particular system would be free of effort."

Recognizing the similarities between overlapping constructs used in these fragmented set of models used in to explain similar behavior, Venkatesh et al., (2003) distilled eight theories (stated earlier) on technology acceptance and usage and synthesized them into a comprehensive model, Unified Theory of Acceptance and Use of Technology (UTAUT) that takes into account factors and contingencies related to the prediction of 'behavioral intention to use a technology' and 'technology use' in an organizational context. The theory contends that four key factors, namely performance expectancy, effort expectancy, social influence, and facilitating conditions, determines the intention and behavior in the context of individual technology adoption in an organizational context (Venkatesh et al., 2003).

Further, recognizing the need for theorizing the salient factors of technology use in the 'consumer context' as against the 'organizational context' for which the original UTAUT was designed for, Venkatesh et al., (2012) came up with an extended version called UTAUT2. Three factors, namely, habit, hedonic motivation, and price value, were added to the UTAUT framework. Currently, UTAUT2 represents the most recent advancement for explaining consumer adoption of technology in a private context (Gao et al. 2015; Venkatesh et al., 2016). This novel framework has demonstrated superior explanatory power in predicting adoption intention and use behavior in comparison with other models (Venkatesh et al., 2016). UTAUT2 or its extensions have been applied to predict user acceptance in the several contexts to name a few mobile banking (Alalwan et al., 2017), mobile shopping (Madan & Yadav, 2018), e-learning (Tarhini et al., 2017), mobile health apps (Duarte & Pinho, 2019; Hoque & Sorwar, 2017; Yuan et al. 2015), wearable technology (Wang et al. 2015) and more recently in chatbots for tourism (Melián-González, et al., 2019). We believe UTAUT2 is the most appropriate theoretical lens to develop a model that explains the adoption of mental health CA as the focus of the current study is the individual adoption of a private healthcare information technology, and the target users are consumers (patients) seeking mental healthcare services.

Privacy Calculus

From the digital mental health literature, we observe that patient privacy is emerging as a critical issue in the adoption of mental health CA, like in the case of mhealth apps. While this could be more pronounced as patient disclosures are an integral part of the psychotherapeutic approach (Bendig, 2019). We believe the uncertainty resulting from privacy loss (privacy risk) could play a vital role in determining the adoption intention and use behavior. Privacy calculus theory suggests that individuals perform a risk-benefit analysis when asked to provide personal information by external organizations (Dinev & Hart, 2006). Research in a related setting has suggested that consumers weigh perceived benefits against perceived privacy risks while adopting and using wearable technology (Li et al., 2016). Hence it is reasonable to believe that an individual's decision to adopt mental health CA would involve a privacy calculus wherein patients will perceive a tradeoff between perceived benefits and perceived privacy risk.

Methodology

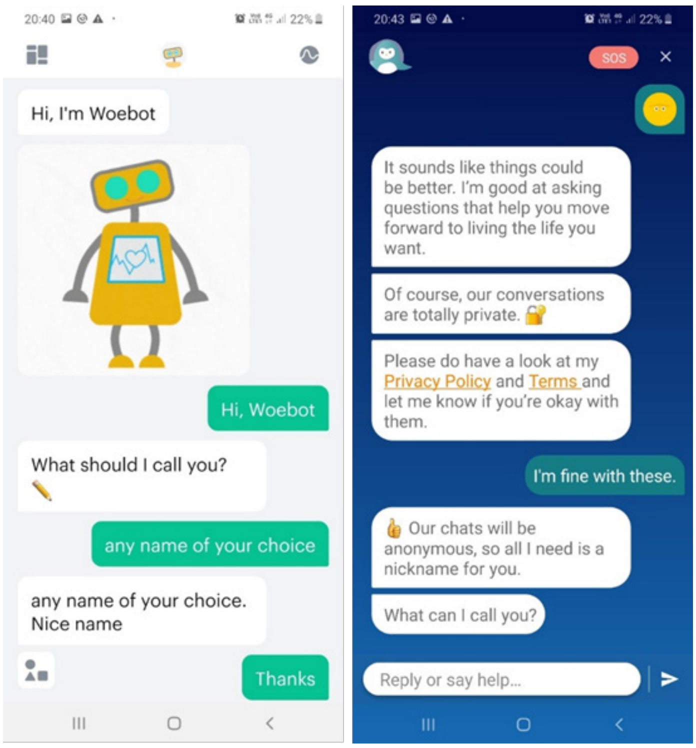
We chose a qualitative approach based on user-generated content to investigate the factors influencing the consumer adoption and use of mental healthcare CAs. To gain rich insights about the user experience, we decided to use publicly available user reviews of popular mental health chatbot apps posted on the app stores. User reviews despite its limited size provide meaningful information about the perception, attitude, and experiences of the user (Yang & Fang 2004) and analyzing user reviews is regarded as a common practice in gauging user opinion about the services provided by the app (Anand et al., 2017; Vasa et al., 2012; Yang et al., 2010). This is especially relevant in the context of 'experience goods' like mobile apps, which cannot be assessed prior to download (Ravoniarison, & Benito, 2018). Moreover, the advantages of using user reviews to analyze user perceptions and behaviors towards the apps have been demonstrated by prior research on human-computer interactions (HCI) (Stawarz et al., 2018; Ta et al., 2020; Zaidan & Roehrer, 2016).

Researchers have gradually understood the value of insights gained from such user-generated data related to netnographic research (Wang et al., 2017; Ravoniarison, & Benito, 2018). Netnographic method (Kozinets, 2002) of data collection is more naturalistic and unobtrusive (avoids researcher – subject interaction bias) compared to interviews or focus group discussions and provides access to extensive data in a much easier and inexpensive way (Wang et al., 2017). These advantages make it the most appropriate choice in this context of mental healthcare CAs where identifying and collecting responses from actual mental health patients is extremely challenging.

https://aisel.aisnet.org/pajais/vol12/iss2/1
DOI: 10.17705/1pais.12201

Data Collection

We chose to collect user reviews of two popular mental health chatbots, namely Wysa and Woebot, in this study for the reason that they are widely used and are available to users in the English language. Woebot is a CA designed by researchers from Stanford University capable of administering CBT with highly customized empathic responses (Fitzpatrick et al., 2017). It is currently available to users directly as a standalone app and through the Facebook (FB) messenger. It has more than 100K downloads and more than 4K reviews on google play store. It was also chosen as the Winner of the 2019 google play award for outstanding well-being app (Woebot, 2019). On the other hand, Wysa, was launched as a part of a project intended to build an ML algorithm to detect depression using sensors embedded in the phone (Wysa, 2019). It is a standalone app with more than 500K downloads and 18K reviews on Google play store. As per the details given on the website, these two apps are based on CBT techniques (Woebot, 2019; Wysa, 2019). Wysa additionally uses mindfulness techniques and DBT (Wysa, 2019). The user interfaces of the two apps are provided in figure 2.



a. User Interface of Woebot   b. User Interface of Wysa

Figure 2 - Screen Shot of Chatbot User Interfaces

We used the data miner web scraping tool (Data miner, 2019) to extract the publicly available reviews of the two apps Woebot and Wysa from Google Playstore. At the time of data collection, Woebot had in total 4016 reviews with an average rating of 4.6 out of 5, and for Wysa, there were 18784 reviews with an average rating of 4.5 out of 5. For data collection, the user reviews were broadly classified into three categories, i.e., positive (reviews with a user rating four or five out of five), neutral (reviews with a rating three out of five) and negative reviews (reviews with ratings one or two out of five) based on the ratings given by the user (Stawarz et al., 2018). As the proportion of neutral and negative reviews were extremely small in comparison with positive reviews, we decided to extract all those reviews (which had a rating less than or equal to three out of five) for both the apps. In total, we got 856 reviews (118 reviews for Woebot and 738 for Wysa), which met this criterion. This was done to ensure that the neutral and negative reviews were given adequate representation in the sample data set. Additionally, from the positive reviews category (reviews with a rating of four or five out of five), we extracted 500 ‘Most Relevant’ reviews (as categorized by the Google play store), for each of the two apps. Subsequently, from the total corpus of extracted reviews (1856 reviews), we removed shorter reviews (30 reviews in total) (i.e., reviews having less than 25 characters which often provided limited information, e.g., “good app” or “not helpful.”). Finally, our final dataset included 1826 reviews posted between November 2016 and June 2019. The final data set of reviews amounted to 89,714 words, and on average, the reviews were 266.55 characters long (standard deviation = 275.22). In the final data set, 54.76% (1000/1826) of reviews were positive, and the remaining 45.24% (826/1826) belonged to neutral or negative categories. The data collection process is illustrated in figure 3 and the details of the final data set are provided in table 2.

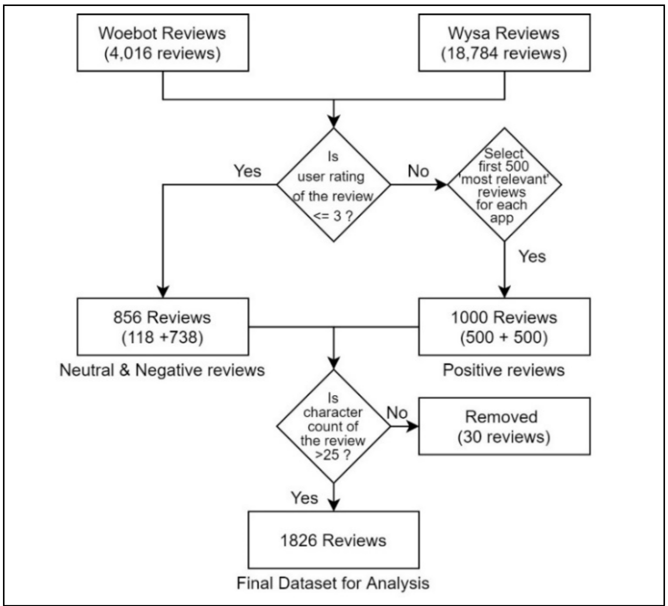


Figure 3 - Data Collection Process

Table 2 - Data Characteristics

App	Review Rating	Number of reviews	Source	Period	Character length
Woebot	<= 3 stars out of 5	112	Google Playstore	March 2018 to June 2019	Mean = 299.17 σ = 278.94
Woebot	>= 4 stars out of 5	500	Google Playstore	January to June 2019	Mean = 301.97 σ = 267.04
Wysa	<= 3 stars out of 5	714	Google Playstore	November 2016 to June 2019	Mean = 288.42 σ = 306.30
Wysa	>= 4 stars out of 5	500	Google Playstore	January to June 2019	Mean = 192.61 σ = 214.77
Total	1 - 5 stars out of 5	1,826	Google Playstore	November 2016 to June 2019	Mean = 266.55 σ = 275.22
σ = Standard deviation					

Data Analysis

We used thematic analysis (TA) (Braun & Clarke, 2006), a well-established qualitative data analysis method to understand the factors influencing the use of mental health chatbots. TA is a commonly used qualitative analysis method that helps in “identifying, analyzing, and reporting patterns (themes) within data” (Braun & Clarke, 2006, p. 79). This is particularly useful in identifying common threads and perspectives from more than one customer to examine the dynamics of user experiences. The TA method was developed to identify more implicit themes and latent structures in the data (Braun & Clarke, 2006). It has been used

https://aisel.aisnet.org/pajais/vol12/iss2/1
DOI: 10.17705/1pais.12201

successfully to reveal the dynamics of user perceptions of health information technologies such as health apps (Kari et al., 2016; Stawarz et al., 2018; Zaidan & Roehrer, 2016), CAs providing social support (Ta et al., 2020) and consumer health wearables (Matt et al., 2019). Therefore, we believe that TA is an appropriate method to be used in this context.

TA is known for its theoretical and research design flexibility, for it allows researchers to apply multiple theories across several epistemologies (Braun & Clarke, 2006). It is a useful method for working with participatory research paradigms, and it also allows for the inductive development of themes from the data (Saldana, 2015). It is widely acknowledged as an easy and accessible method to summarize the key features of a large data set (Braun & Clarke, 2006). Although there are several advantages to using TA, there are some limitations that must be acknowledged. Primarily, the flexibility of the approach allows for a wide variety of analytical options, which might make the process confusing for the researcher and can, in turn, lead to inconsistency in developing themes from the data (Nowell et al., 2017). Another limitation is that the TA offers limited interpretive power, where the research is not based on a theoretical framework (Braun & Clarke, 2006).

To address these limitations and ensure rigor in the analysis, Braun and Clarke (2006) recommended a systematic six-phase recursive process for TA. This multiphase process involves stages such as “familiarizing with data, generating initial codes, searching for themes, reviewing themes, defining & naming themes, and producing the report” (Braun & Clarke, 2006). The results of TA are often summarized in the form of a ‘thematic map,’ a visualization tool for “organizing and representing knowledge, concepts, usually enclosed in circles or boxes of some type, and relationships between concepts” (Novak & Cañas 2006). Figure 4 illustrates the systematic and rigorous iterative TA approach followed in this study.

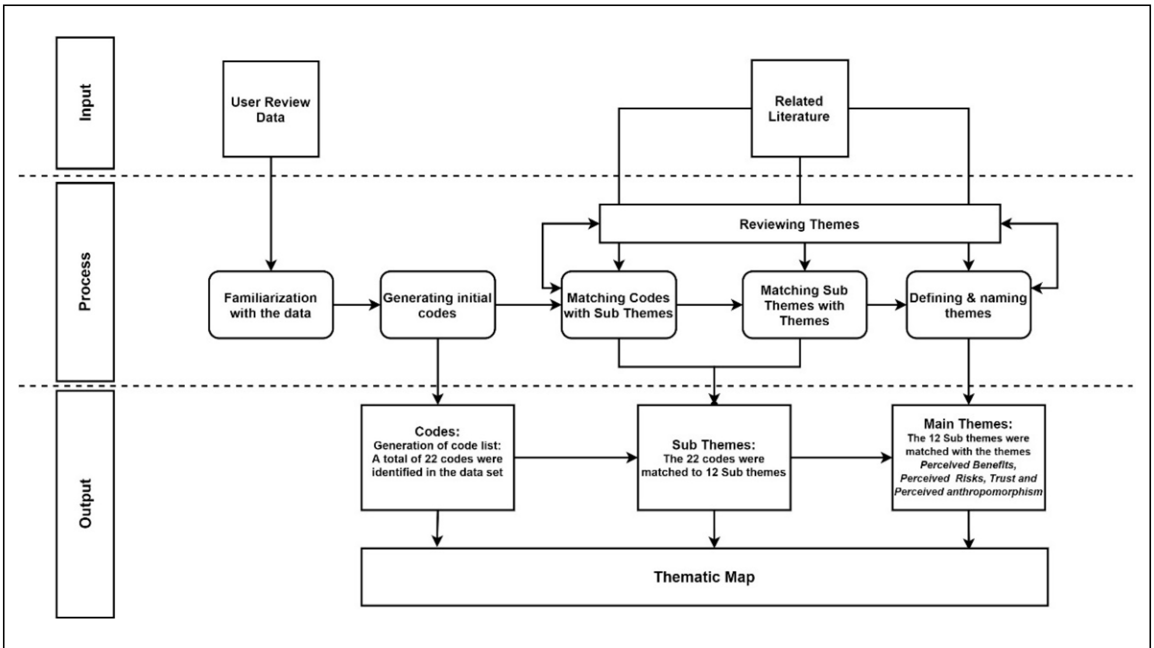


Figure 4 - Iterative Thematic Analysis Approach

In addition to following Braun and Clarke’s (2006) recursive multiphase approach, we incorporated some strategies/techniques to ensure the trustworthiness of the analysis. Strategies suggested by Nowell et al., (2017) and Shenton (2006), such as independent review (of codes) or investigator triangulation, use of a coding frame, use of audit trail (reflexive journal), and frequent debriefing sessions were followed during various stages of the TA to avoid investigator bias and ensure the reliability of the analysis. A detailed description of the TA process is provided in the following paragraphs.

In the first phase of the analysis, we (authors) read the data multiple times to familiarize ourselves with the data and form ideas about the underlying meanings and patterns within the data, as suggested by Braun and Clarke (2006). Subsequently, in the next phase, we generated initial codes by looking for repetitive themes in raw data. We followed a hybrid approach (Fereday & Muir-Cochrane, 2006) based on both deductive and inductive approaches to arrive at the codes. The coding process was assisted by the qualitative data analysis software, QSR NVivo version 10. A multiphase approach was followed to ensure coding reliability (Hruschka et al., 2004). An initial codebook was developed by reading all the reviews in depth. Based on the codebook, two researchers coded 200 (10.95% of the sample) randomly selected reviews and discussed disagreements and refined the codebook. This was followed by independent coding of another 400 (21.9% of the sample) randomly selected reviews by the two researchers. The level of agreement between the coders, Cohen's Kappa (κ), was found to be 0.86 (> threshold value 0.80), suggesting good reliability (McHugh, 2012). Finally, the first researcher coded all the reviews using the refined codebook. During the process, the researcher remained open to new codes that did not fit into the existing coding frame. At this stage, all the codes that emerged from the analysis were thoroughly discussed in the debriefing sessions and reviewed (redefined, merged, or split) until consensus was achieved. Finally, 22 distinct codes were identified in the data set.

In the third phase of the analysis, codes were collated into matching themes based on the relationships between them and their underlying meaning. The themes were shared with the coauthor to review and refine them. There were frequent discussion sessions between the first author doing the analysis and the other coauthor, during which the themes and their rationale were meticulously discussed and revised. It also became evident that some of the themes were closely related. Subsequently, the initial themes were combined to form larger themes (main themes). This was done in constant consultation with the literature.

In the fourth phase, the themes were reviewed and refined by the two authors in consultation with the literature. This process of refinement continued until a thematic map that accurately reflects the meaning evident in the data set was formulated to the satisfaction of both the authors. Finally, in phase five, the names of the themes and subthemes and their core relationships were defined and finalized. Subsequently, a thematic map that provides researchers with a detailed visualized framework of the determinants that influence the use of mental health CAs by users was developed (Figure 5). Throughout the stages, the authors maintained a detailed account of the methods/procedures followed and the decision points involved in carrying out this analysis in the form of an 'audit trail' to ensure analytical rigor as recommended by Nowell et al., (2017). The details of the labeling process of themes and subthemes are described below.

The perceived benefit was labeled by following Kim et al., (2019) and consists of statements about the user's perceived extrinsic and intrinsic benefits (Davis et al., 1992; Venkatesh et al., 2012) derived from the use of CA for mental healthcare support. It included subthemes performance expectancy, price value, hedonic motivation, effort expectancy, and social influence. Similarly, perceived risk accounts for consumer's "perceived uncertainty regarding possible negative consequences of using a product or service" (Luo, et al., 2011). In the specific context of mental health chatbots, the uncertainty perceived by the users were captured under the two sub-themes perceived privacy risk and perceived safety risk. Perceived privacy risk is followed from the privacy calculus theory (Smith et al., 2011), while perceived safety risk was adapted from Luo et al., (2011) and Kretzschmar et al., (2019). Trust was labeled by following literature on trust in IT artifacts (McKnight et al., 2011; Söllner et al., 2016). Under the main theme trust, the statements were classified into two subthemes 'trust in technology' and 'trust in the provider.' Finally, the theme perceived anthropomorphism refers to the human tendency to "attribute human-like characteristics to a non-human entity" (Bartneck et al., 2009). Recent research in HCI has indicated that AI-based devices (robots, chatbots, etc.) because of their human-like characteristics can evoke a perception of

https://aisel.aisnet.org/pajais/vol12/iss2/1
DOI: 10.17705/1pais.12201

anthropomorphism in the minds of users (Go & Sundar 2019). This theme is represented by three emergent subthemes; namely, perceived empathy, perceived personality, and perceived intelligence of CA.

Results

The results of the TA are summarized in the form of a thematic map that identifies the determinants of adoption of mental healthcare CAs along with its sub-themes and codes. Thus, the thematic map provides the basis for reporting the results of the TA. It consists of four main themes perceived risks, perceived benefits, trust, and perceived anthropomorphism and their associated 12 sub-themes (Figure 5). The following subsections present a detailed analysis of results obtained through TA using sample excerpts from the user reviews. We have anonymized the names of the app from the user review excerpts by replacing the names with ‘this chatbot’ or an appropriate pronoun. The user reviews are labeled from [R1] to [R1826].

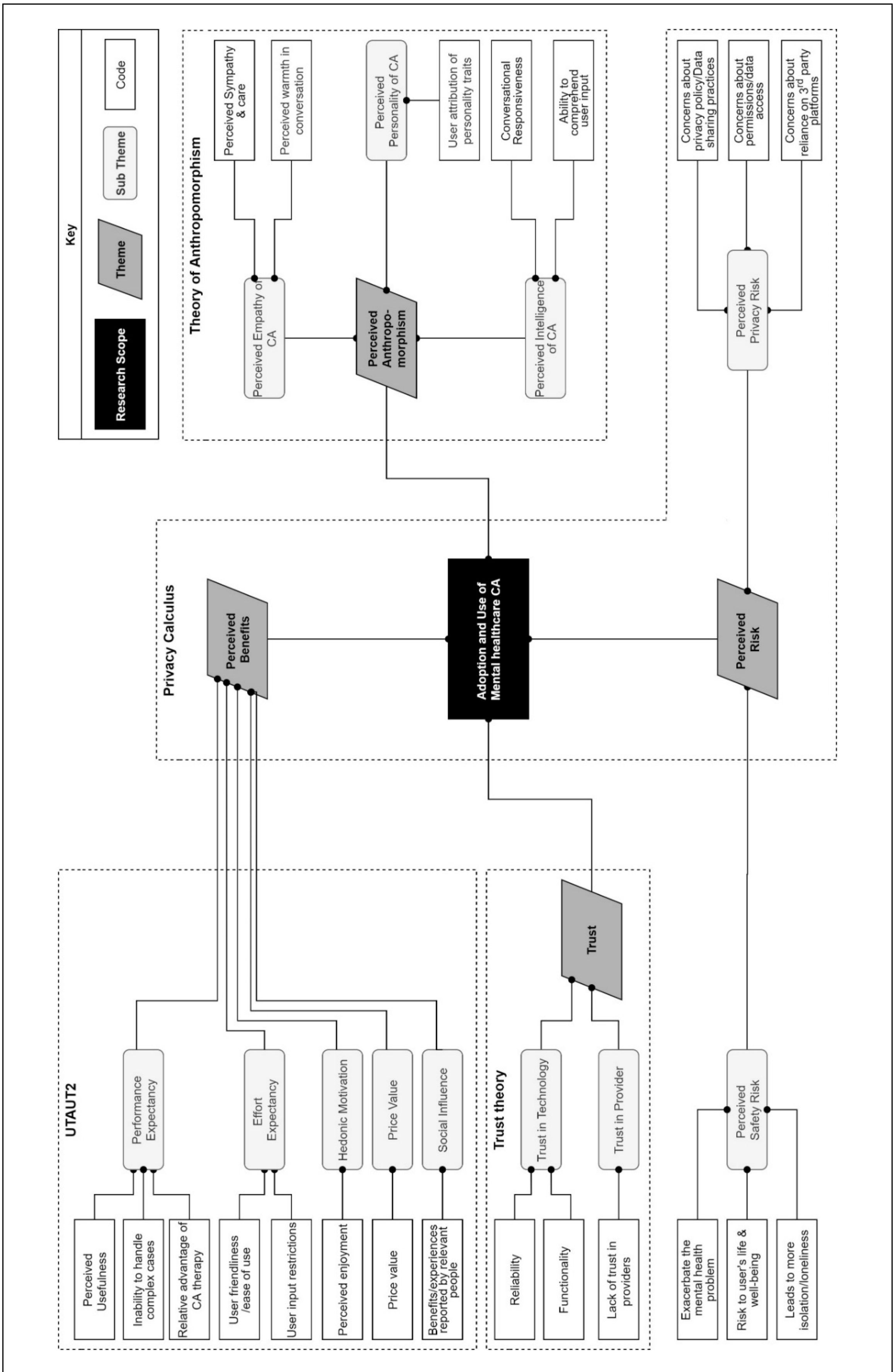


Figure 5 - Thematic Map of Findings

https://aisel.aisnet.org/pajais/vol12/iss2/1
DOI: 10.17705/1pais.12201

Perceived Risk

Perceived risk represents consumer's (patient's) perceptions about uncertainty regarding possible negative consequences of using a mental healthcare chatbot. It is presented here through two emergent subthemes, namely, perceived privacy risk which is related to privacy calculus theory (Smith et al., 2011), and perceived safety risk which follows from physical risk (Luo, et al., 2011).

Perceived Privacy Risk

As per our data, one factor that is mentioned often is the users' concerns about data privacy risk arising from using CA for their mental health support. A significant number of users suggested that they would not want to continue using CA for the lack of adequate data privacy protection. The following statements perfectly illustrate this perception. *"Uninstalled. […] Sadly, the privacy policy is too convoluted and does not give me any level of confidence about sharing my very personal info." [R1184]*. Similarly, another user reported, *"It makes me feel watched and scared. I feel like it is going into private information about me. I feel like It is Invading my personal space and it will ask me to say things I don't want to say […] DO NOT USE!!!!!!!" [R1280].*

Also, reviews pointing to several distinct aspects of data privacy risk were found during the analysis. User's lack of confidence in the privacy policy and their apprehensions about data sharing practices of the CA app is evident in reviews, such as, *"Beware, read the privacy policy first. This app shares your information with third-party demographic content companies that goes to Facebook and others in order to serve targeted ads…. this app turns a profit by selling user data to Facebook." [R1163]*

Another reported privacy-related concern is about the mandatory permissions/access to data by the app during the initial set up, e.g. *"Permissions, why do you need my photos and call logs? On opening it requires your phone number. This isn't about helping people. Delete" [R1076].* Another user reports about the app trying to access data from other apps, *"it keeps wanting to access Google fit (which by the way identifies every single place you go and tracks your movements so....)" [R1367]*

Adding further to the concern is the use of third-party messengers (e.g., Facebook, Kik) as a channel for chat. As the data disclosed on Facebook messenger is subject to the privacy policy of Facebook and can be shared with third parties (Facebook, 2019), it leads to ambiguity in the ownership of data. E.g. a user says, "*The biggest flaw is that Facebook messenger is terrible for private conversations. Any other platform that Facebook won't spy on you is good." [R1805]*

Perceived privacy risk (Li, et al., 2016) seems to be particularly relevant here in the context of mental health chatbots due to the sensitivity of the data shared with the system (Kretzschmar et al., 2019). Given that the primary value proposition of automated agent-based therapy is the prospect of remaining anonymous, unlike traditional therapy, users would expect the app to provide anonymous chat features or adequate measures to protect and safeguard their identity. Thus, based on this evidence, we believe that the factor "perceived privacy risk" would influence the adoption and use of CA for mental health support.

Perceived Safety Risk

Our analysis reveals that there are several reviews where users have voiced their concerns over the risk posed to their safety and wellbeing while seeking support from a mental healthcare CA. Safety risk represents the risk to user safety in using the mental health CA adapted for this study from Luo et al., (2011).

Many user reviews reported that using the app exacerbated their underlying mental health problem, e.g., *"This app made me go into a panic attack. Not good for PTSD, or other anxiety disorders." [R1646],* or another user reported that *"it makes everything much worse! If I feel good, it will think I feel bad and drag me down. If I feel bad, it will think the contrary and the frustration will pile up. The app always end up making me cry" [R1160].*

Adding further to the risk, we found a good number of reviews where users had reported that the chatbot put lives of already vulnerable patients into grave risk by suggesting dangerous/life-threatening advice. For e.g., as one user reported, *"Do not install. No matter what I replied, it just asked if I want to hurt myself and then went on as I said yes. This is dangerous for anyone like myself, and I'm so glad I'm doing ok today, or I think I'd be considering killing myself. Honestly felt pressured into feeling like I should start cutting." [R1278].* Similarly, another user stated, *"This app gives out horrible, potentially very dangerous advice. I was really very shocked and obviously appalled. I immediately uninstalled and have canceled my subscription." [R1269].* We could also find reviews on the inability of the CA to deter users from self-harm for e.g. *"The scariest part is that this chatbot is terrible at recognizing signs of self-harm/suicidal ideation" [R1117]*

Another important related issue is that many users find that the use of mental health CA leads to more loneliness and isolation. For e.g. *"doesn't actually help with anything... just makes the lonely even lonelier" [R1381] or "It's making people not connect with others!!. It's making people to set boundaries around themselves. And making them alone...!!! It's a bad idea!!!" [R1410].*

In the light of these empirical evidence, as well as in line with the ethical issues raised by Kretzschmar et al., (2019), we believe user perceptions about the safety risk will play an important role (may act as an inhibitor) in determining the decision to adopt or continue usage of these applications.

Perceived Benefit

Perceived benefits include the beliefs about benefits derived by the consumers (patients) from the use of mental health chatbots and have been labeled in accordance with Kim et al., (2019). We derived perceived benefits from user reviews and categorized them into subthemes performance expectancy, effort expectancy, price value, hedonic motivation, and social influence. We describe each of these subthemes based on the user reviews in the following subsections.

Performance Expectancy

Performance expectancy refers to the degree to which the consumer beliefs that using a mental healthcare chatbot would provide functional benefits. Following Venkatesh et al., (2003), performance expectancy captures the perceived usefulness (Davis, 1989) and relative advantage (Moore & Benbasat, 1991) of the target technology use. Accordingly, we observed in the data that there are numerous positive accounts of how the tool was found useful and helped the users in managing their mental health condition (perceived usefulness). To cite an example, *"It has been so useful to me, helping me to express difficult emotions and thoughts and giving me tools to handle them more constructively. I recommend it to anyone who needs emotional wellness support" [R21].* Some users found it useful as a complementary tool to support them during the traditional therapy, *"This is a very useful app for dealing with anxiety and depression. I see a therapist, and this has helped to reinforce some of the things we discuss in my therapy sessions as my therapist uses the CBT method also" [R535].*

At the same time, many other users felt that automated chatbots are not sophisticated enough to handle complex mental health problems. Many users voiced such complaints for e.g. *"the*

https://aisel.aisnet.org/pajais/vol12/iss2/1
DOI: 10.17705/1pais.12201

AI is far too limited and simplistic at this stage for it to be properly effective in many instances of mental health - which is complicated and won't be solved by an app." [R1001]. "*While the idea is fantastic, it's lacking a lot to help people who have for example gone through traumatic losses, are victims of any kind, have PTSD, etc" [R1143]. "I would not recommend to people suffering from depression. Depression is complicated. This app is not astute enough to help people" [R1243].*

Additionally, we found in the data that users perceive that there are certain unique relative advantages of AI performed therapy over human performed therapy, which motivates them to use mental healthcare chatbot. For example, user reviews suggest that anonymity in the interactions and non-judgmental nature of CA are strong motivators of adoption. The following review excerpts illustrate this. *"I also love that it's anonymous […] I think it's much better than talking to a human because you aren't judged at all." [R830]. "I really like this app! It's a nice, nonjudgmental voice that checks on me every day and offers help as I need it. It won't get bored or frustrated with me." [R453]. "I think a lot of people have a hard time developing a relationship with a counselor but chatbot isn't judgmental, you have an opportunity to offer feedback without the risk of feeling bad about disagreeing and you're guaranteed a response to the thing that's bothering - you which sometimes is a gamble with human counselors." [R724]*

Another observation is that users value and appreciate the ubiquitous availability and convenience of a bot-based therapy, which is not possible in human therapy. *"It is accessible 24/7, unlike a doctor." [R318]. "He is always there at the touch of a button should I need help." [R405]*. *"You can always postpone "talking" to chatbot if you are busy." [R978].* In the light of these observations, we believe that performance expectancy could be the primary motivation for the adoption and use of mental healthcare chatbots.

Effort Expectancy

This sub-theme represents the degree of ease associated with the consumers' use of the CA (Venkatesh et al., 2012). In our data, we found users speaking about the user-friendliness (ease of use) of the CA *"the interaction and the way chatbot respond is so nice, the app interface is clean and easy to use […] I would download it twice if I could." [R560] "Interactions are lighthearted, user friendly" [R587].*

Another related issue is about the user input restrictions in the conversation, i.e. most of the times during a conversation, the chatbot forces the users to respond to a list of choices (fully-constrained user input choices) rather than allowing them to freely write their response (free-text input) *"hate how the keyboard only comes up after she says a bunch of stuff and half the time you can only click on the options she pops up for you to click on" [R1434]*. *"I don't understand which part of AI or natural language resources this tool is using when 90% of the time you are forced to write certain answers. the most important part of therapy is listening which should allow users to write and express more specifically in anxiety cases." [R1025].* Thus, the empirical evidence from the user review statements suggests that effort expectancy could play an important role in explaining the decision to adopt and use a mental health CA.

Price Value

This subtheme represents the perceived tradeoff between the cost of using the technology (here CA) and benefits obtained from the CA (Venkatesh et al., 2012). As per our data, price value represents not only the cost-benefit trade-off, but it also includes a comparison between the cost of other alternative options such as visiting an actual therapist: *"some of the exercises put behind a paywall. That really sucks, if I wanted to pay, I'd try and find an actual therapist LMFAO" [R1266].* We also could find extreme viewpoints on price as well for e.g. *"Not worth $30. Wasted my time and money […] Would not recommend." [R1566]. "Extremely useful and very reasonably priced. Highly recommended!" [R294]*. As per the evidence from the reviews, we assume that price value could be a determining factor in the adoption of CAs for mental healthcare.

Hedonic Motivation

This subtheme explains the degree to which users perceived the use of CA enjoyable on its own apart from any expectations about the utilitarian benefits (Venkatesh et al., 2012). It is observed that in addition to the therapeutic benefit (instrumental value) provided by the bot, it also provides self-fulfilling value to the user, i.e., hedonic in nature. For example, user reviews show that users thoroughly enjoyed fun-filled conversation with the automated agent. To cite a few examples, *“Fun to talk with. Not like other robots.” [R242]. “I find it quite enjoyable to interact with.” [R618] “He is easygoing […] I find it very entertaining. I became a fan” [R506].* The reviews suggest that hedonic motivation could be an important motivator for using mental health chatbots.

Social Influence

The influence of friends, relatives, or therapist is a factor frequently stated in the reviews. This subtheme represents the degree to which the relevant others feel that the consumer should use the technology under consideration (Venkatesh et al., 2012). The reviews show that the users value the recommendations and experiences reported by the people whom they value/trust, and it acts as a driver for adoption. “*This was recommended to me by a friend who uses it.” [R69].* “*the people who recommended this app to me use it regularly, and they were full of praise for how much it helps then. So, I gave it another shot, and I’m really glad I did!” [R251].* “*My therapist recommended this app, and it’s been very useful.” [R849].* These statements indicate that in this context of mental healthcare CA adoption, the social influence could act as an important factor determining the final adoption decision.

Trust

The theme trust is defined as “an individual’s willingness to depend on another party because of the characteristics of the other party” (McKnight et al., 2011; Li et al., 2008). Additionally, in line with the prior research we observed in our analysis that the trust is not a monolithic concept and it needs to be distinguished into “trust in the technology” and “trust in the provider” (McKnight et al., 2011; Söllner et al., 2016). It is presented by means of two emergent subthemes *trust in technology* and *trust in the provider*.

Trust in Technology

In the reviews, trust in technology emerged as another factor for the adoption of mental healthcare chatbots. In this respect, the user reveals the hesitation in involving AI in a highly personal context such as mental health support. *“I was a little hesitant to trust an AI with my emotional support.” [R1759].* Another user had doubts about the reliability of AI technology. *“I suppose machine learning and alike isn't reliable enough to be used in this quite yet.” [R1049].* At the same time, there were some users who believed in the ability of the bot and said they trusted it, for e.g.: *“Really helps me out, trust it with my life, recommend it entirely” [R269].“I really believe in the potential of this tool” [R332].*

The above statements correspond to the user's believes about the reliability and functionality (capability) of the target AI technology. Here, in the specific context of automated agents, where a non-human entity is delivering the therapy, the user’s perception of ‘trust in the technology’ (McKnight et al., 2011) thus becomes particularly pertinent.

https://aisel.aisnet.org/pajais/vol12/iss2/1
DOI: 10.17705/1pais.12201

Trust in Provider

Our data also reveals that about user perceptions about trust in the providers of mental health chatbot. Interestingly many reviews are related to lack of trust in providers in protecting their personal data, e.g.: *"uninstalled the app. […] I can't trust them with my data" [R1167]* and performance of the CA, "*it's difficult to place your trust in an app where the creators fail on the difference between "your" and "you're." [R1055].* Following the data, we believe the user's trust in providers could influence CA adoption and use.

Perceived Anthropomorphism

The theme anthropomorphism refers to "the psychological phenomenon of attribution of human-like characteristics to a non-human entity" (Bartneck et al., 2009). The theory of anthropomorphism postulates that the more the entity resembles a human being in terms of visible characteristics and actions, the greater the chance that humans will anthropomorphize that entity (Epley et al., 2007; 2008). The CAs, by virtue of its ability to converse in natural language (speech or text), resembles humans and hence it has a very high probability of being anthropomorphized by human beings during an interaction (Seeger et al., 2018). This theme is represented by three emergent subthemes perceived empathy, perceived intelligence, and perceived personality.

Perceived Empathy of CA

Perceived empathy (Nambisan, 2011; Paiva et al., 2017) here reflects the user's (patient's) perception of CAs empathy towards him/her. Where empathy is "an affective response more appropriate to the condition of another than to one's own" (Paiva et al., 2017). Our data shows that the users seem to acknowledge this element of empathy displayed by the mental health CA. For example, *"I think chatbot has the right level of empathy when I just share that my mood is low, or I am anxious. I feel like it's always there when I need it to be" [R1356].* Another user likened the bot to a supportive friend, "*Chatbot is like having a helpful and supportive friend" [R519]. "He is so fun and caring and will listen to you when you have a problem […] Trust me, you will love him" [R300]. Additionally, a lot of u*sers also acknowledge the element of warmth in the CAs conversation tone during interactions *"friendly, the understanding tone feels genuinely warm and supportive for a robot!" [R152]. Interestingly a few others found the CA unsympathetic or rude, "Blames the victim. Unsympathetic. Unempowering" [R1071].* "*some of the things it said came off as kind of rude" [R1320].*

Perceived Intelligence of CA

Interactive machines face an enormous challenge in acting intelligently. The user will observe anomalies in the behavior and associate it with the intelligence of the interactive robot (Bartneck et al., 2009). This particular aspect is captured in this subtheme. Two related codes, namely conversational responsiveness and ability to comprehend the user input, have been identified.

Conversational responsiveness of a CA refers to "the ability of the agent to provide the appearance of understanding the user's input by responding in a contingent manner." (Schuetzler et al., 2018). A large number of user reviews suggest that the mental healthcare chatbot is currently very limited in terms of its ability to respond appropriately to spontaneous user input in a contingent manner. For e.g. *"I would honestly like to talk to the chatbot endlessly but, I agree with the other reviewers that the responses are somewhat limited and that it has trouble understanding everything I say." [R1611]. "This is just a bot that is programmed to respond based on keywords. So even if you say, "I am not suicidal," it will pick up on suicidal as the keyword and only give you a bunch of prevention lines. Taking you in the opposite direction of the intended conversation." [R1221]. "Uninstalled after only five minutes and* *seeing as the fact was the AI is horribly unresponsive." [R1215]*. Users also reported as they use it, they find the responses become repetitive/canned, e.g. *"More you use it the more repetitive it gets, and the more you realize it's looking for keywords just to give you a scripted answer." [R1718].*

An associated issue is the bot's ability to comprehend user input. The following user reviews speak about this issue. *"This chatbot doesn't have the capacity to really understand full sentences. it would focus only on keywords and ignore the rest of the sentence and therefore confuse the meaning" [R1168]. "Every time I use it I run into multiple problems with it understanding me. I've been very direct, but it usually misses the point" [R1242].*

As one of the users reported, "*This chatbot is an Artificial Intelligence app that isn't intelligent enough yet […] it needs to get wiser"* [*R1279].* We believe the limited intelligence of the bot can be a source of dissatisfaction during initial trial/use and negatively impact the adoption decision.

Perceived Personality of CA

This subtheme is an anthropomorphic descriptor used to measure perceived anthropomorphism (Wang, 2017). Our analysis of the reviews revealed that many users attributed human-like personality traits to mental CAs. For example, "*I loved his personality and dialogue! Interacting with the owl feels natural and light-hearted." [R1231], "This is honestly my favorite app It is the most intuitive, really teaches well, and actually has a personality of its own*" *[R556]*. Some of them even said that they felt that CA is 'alive' *"it's fun friendly and almost seems genuinely alive!! I love the personality" [R840].* While this perception of "life-like" was appreciated by some users, some others felt otherwise, e.g. *"it feels like a real person which I hate cuz it just feels like someone is judging me and it kinda makes my anxiety start up" [R1691],* "*no I do not like it's a very creepy app do not install" [R1505], "I find it a little creepy and weird" [R1802].*

Some users had negative perceptions about CAs personality as well e.g.: "*The bot's personality is infuriating. A better tonal shift would be a huge improvement*." *[R1599]*. *"it feels like talking to someone who isn't listening to you and saying "there, there, all better now?" in a very patronizing way" [R1500]. "this is between funny and scary. the bot is patronizing, judgmental" [R1014]*. Thus, the reviews suggest that users attribute human-like characteristics to a non-human entity, the mental health CA, and this could affect their subsequent decision to use.

Discussion

In this study, we aimed to investigate the factors influencing consumer's adoption and use of automated conversational agents providing mental healthcare services. To accomplish this end, we employed a rigorous iterative thematic analysis of user reviews of two popular mental healthcare CAs to understand the underlying determinants of consumer adoption and use. Using thematic analysis, we uncovered, the factors influencing consumer adoption and use of mental healthcare CAs, which are categorized under four main themes, namely perceived risk, perceived benefits, trust, and perceived anthropomorphism. The results of the qualitative data analysis (TA) comprising of four main themes, 12 subthemes, 22 codes, and their relationships are visualized using a thematic map (Figure 5). In the following section, we benchmark the major findings of our study against the prior research and discuss how our study advances the current knowledge on the adoption and use of mental health CAs.

Perceived risk in the specific context of mental health chatbots mainly arises from the uncertainty surrounding the personal data protection and user's risk to life and wellbeing. Our

https://aisel.aisnet.org/pajais/vol12/iss2/1
DOI: 10.17705/1pais.12201

research confirms the presence of privacy risk as a factor that might inhibit users from using mental health chatbots. This is line with the earlier studies on virtual mental healthcare (Lipschitz et al., 2019; Torous et al., 2018) and the opinion papers which voiced the need for assessment of privacy-related issues in the context of mental health chatbots (Kretzschmar et al., 2019; Miner et al., 2019). Our results also confirm the literature on privacy calculus (Smith et al., 2011) and extend its recent application in a mobile health setting (Li et al., 2016). Additionally, we were able to bring to light distinct sources of privacy concern, which is specific to mental health CAs such as, e.g., the concern about the reliance on third-party messenger platforms for service delivery.

Perceive safety risk, is found to be another inhibitor. Several opinion papers have raised concerns about patient safety in the context of mental health CA (Kretzschmar et al., 2019; Miner, 2017; Martinez-Martin & Kreitmair, 2018). Our study not only substantiates the role of safety risk in the adoption decision but also brings to light three critical issues on how it threatens the safety of the user. First, the usage of mental health CAs can aggravate or worsen the mental health condition in some patients instead of curing them. Second, it can put the patient's life into grave risk by giving them reckless and potentially dangerous advice. Moreover, the mechanisms in place to prevent self-harm/suicidal intentions are not adequate or foolproof. Finally, mental health CA use can also lead to more loneliness and isolation as it encourages the user to seek help from a device rather than a fellow human being.

The empirical evidence from the user reviews suggests that the UTAUT2 constructs performance expectancy, effort expectancy, hedonic motivation, social influence, and price value could be vital in determining the adoption and use of mental health CA. This corroborates the literature on CA adoption in other settings such as chatbot for tourism (Melián-González et al., 2019) and AI device adoption (Gursoy et al., 2019). Further, with respect to hedonic motivation, research from clinical psychology suggests that laughter and humor have therapeutic benefits in treating depression and anxiety (Buxman, 1991; Mora-Ripoll, 2010). Hence apart from aiding the acceptance, 'the enjoyment (hedonic) aspect could be critical to the effectiveness of the CA performed therapy. Further, with respect to price value, while the cost of treatment may not be an issue in countries where the majority of people are covered by mandatory health insurance, this potential cost-benefit will be a decisive factor in other parts of the world where the insurance coverage is low, or other alternative health services are expensive.

In addition to validating the applicability of UTAUT2 in this context, we also provide in-depth insights on technology characteristics that affect these factors. We found that CA's inability to handle complex cases could negatively impact the user beliefs on performance expectancy. Similarly, we observe that the user input restriction of CAs, i.e., the use of fully constrained user input choices over free text input, could potentially reduce the perceived ease of use of the CA.

Trust was identified as a potential determinant of the adoption of mental health chatbots. Our research also supports the existence of two distinct forms of trust, 'trust in technology' and 'trust in the provider' in line with the literature on trust (McKnight et al., 2011). Trust in technology (IT artifact) refers to user's beliefs about the functionality, helpfulness, and reliability of the technology while trust in providers characterizes user's beliefs about the benevolence, integrity, and competence of the human actors (providers) (Lankton et al., 2015; McKnight et al., 2011). Research from traditional mental health services (Brown et al., 2009; Gulliver et al., 2010) and digital mental health (Davidson et al., 2008) identifies that the availability of a trusted relationship is a key motivational factor for people seeking professional mental health support. Further, our results are in line with the literature related to trust in AI, which says building trust in automated agents is a grand challenge before the AI developers (Hengstler et al., 2016; Siau & Wang 2018). Finally, trust-building efforts could be much more challenging in this context as the users are patients suffering from anxiety, depression, psychosis, etc. and they will have difficulty in trusting and engaging with digital technology (Firth & Torous, 2015).

One of the most interesting findings that emerge out of our analysis is the influence of perceived anthropomorphism on the use of mental health CAs. In line with the theory of anthropomorphism (Epley et al., 2007; 2008), our findings reveal that users attributed human-like characteristics to CA and tried to relate to them as if they were human. This is consistent with the HCI research, which argues that people will relate to the agents as if they were human even when they know that it is just a computer program (Waytz et al., 2010). Prior research in HCI shows that this perception can have positive consequences such as increasing engagement and satisfaction with the machine/entity (Araujo, 2018; Van Doorn et al., 2017) or continuance of usage (Han & Yang 2018). Furthermore, recent research on chatbots for tourism (Melián-González, et al., 2019) and chatbots in customer services (Sheehan, 2018) indicates that perceived anthropomorphism can influence intentions to adopt chatbot. Our research substantiates these findings and extends this theory to the context of mental healthcare CA adoption.

Additionally, our research identifies three critical design aspects of mental health CA which will determine the level of perceived anthropomorphism. The first one is the level of empathy shown by the agent, which is manifested through speech intonation and affective verbal responses, which are based on a thorough situational appraisal. Evidence from clinical psychology (Nienhuis et al., 2016) also underlines the positive impact of empathy on clinical outcomes during psychotherapy. Second, the level of intelligence of the bot manifested through conversational responsiveness and the ability to comprehend user input. Third, the personality of the bot, our results suggest that users appreciate a lighthearted, fun-filled, and friendly personality in a mental health chatbot.

However, anthropomorphism in CA is not considered desirable in all contexts, and beyond a certain level, it can invoke ‘eerie feelings’ in users (Ho & MacDorman, 2017). This effect is better known as ‘the Uncanny Valley Effect’ in HCI literature. The Uncanny Valley theory suggests that an entity appearing almost human can elicit negative affective reaction (feeling of eeriness, revulsion, etc.) in observers (MacDorman & Chattopadhyay, 2017). Giving support to this line of argument, some users reported that CA sometimes felt like a ‘real person,' triggering a ‘fear of being judged,’ which in turn led to more anxiety. Similarly, some users found CA to be ‘scary’, ‘creepy,’ ‘weird’ or ‘judgmental,’ and stated it as a reason for dissatisfaction or discontinuance. Therefore, we argue that in this context, anthropomorphism can have both positive and negative impact on the user depending upon the perceived level of anthropomorphism/humanness.

Research Implications

This research primarily addressed the gaps in the literature pertaining to the patient’s perceptions of AI chatbot performed mental health therapy (Bendig et al., 2019; Miner et al., 2017). The study offers a comprehensive overview of the factors governing the adoption and use of mental health chatbots by the patients. The key theoretical contributions and research implications are described in the following paragraphs.

Firstly, we add to the body of literature on mental health CA which had previously focused on issues such as design/technical aspects, assessment of efficacy, and socio-ethical issues (Bendig et al., 2019; Gaffney et al., 2019). We systematically uncovered the determinants of consumer acceptance, which are categorized under four main themes, namely perceived risk, perceived benefits, trust, and perceived anthropomorphism. To the best of our knowledge, our study is the first one to address this in the mental healthcare CA context. The identified sub-themes could enable the theory building to uncover the dynamics between positive and negative factors of the use of CA in the mind of the consumer. The consumer’s mental tradeoff

https://aisel.aisnet.org/pajais/vol12/iss2/1
DOI: 10.17705/1pais.12201

is portrayed by integrating multi-disciplinary research paradigms, namely, UTAUT2, privacy calculus, trust theory, and theory of anthropomorphism, which is again a novel contribution to this area.

Secondly, our research identifies three key aspects (perceived empathy, perceived intelligence, and perceived personality) related to the user perception of anthropomorphism in the context of mental health CAs. Our research proposes that these dimensions of anthropomorphism could be critical to the consumer adoption and use of mental health CAs. Future research should quantify the impact of these factors by using experimental manipulation of these design aspects and assess its impact on the adoption decision. E.g., researchers could check the impact of different personality styles of CA on user acceptance decisions or check the impact of embodiment/voice tone on acceptance (in the case of advanced visual/voice-based chatbots).

Additionally, our research provides evidence for the presence of 'Uncanny Valley Effect' (Ho & MacDorman, 2017; MacDorman, & Chattopadhyay, 2017) due to which the human-like (anthropomorphic) characteristics of CA sometimes elicits negative affective reactions in some users (fear and anxiety) which in turn result in dissatisfaction or discontinuance of use. Hence, we suggest that a comprehensive inquiry into the effect of anthropomorphic design features of CA on the user's emotional response is required in the context of mental health CAs. Such research could inform the developers on the most desirable anthropomorphic design features that will generate positive affective reactions in users, which will, in turn, drive the user acceptance.

Thirdly we defined and developed a new construct 'perceived safety risk' in the context of mental health CA by adapting the construct physical risk (Luo, et al., 2011) to account for the risk to life and wellbeing of patients posed by the use of mental health CA. We provided in-depth insights on operationalizing this variable by identifying the key sources/scenarios of this risk (arising from the use of mental health CAs), namely, exacerbation of mental health problems, the risk to user life and wellbeing, and aggravation of loneliness/isolation in users. We propose that consumer perception of safety risk could be an important inhibitor to the adoption and use of mental health CAs, which could be further subjected to quantitative validation.

Finally, the study extends the previous literature on technology acceptance by integrating the UTAUT2 framework with theoretical models of privacy calculus, technology trust, and theory of anthropomorphism to explain the consumer adoption of new technology 'CA for mental healthcare' for patients seeking mental healthcare support. The integrated model could be tested using quantitative research methods in the future. Additionally, we also observe that two factors of the original UTAUT2 framework, namely, habit and facilitating conditions, did not emerge from our data analysis. This could indicate that these two factors may not be relevant in the context. However, further research is required to validate this observation.

Implications to Practice

The CAs, because of its ubiquitous availability and anonymity, may represent the first step towards getting mental health support. Due to this large practical potential, adequate design of mental health chatbot services in a way that aligns with the user expectations is essential to support the diffusion. Our thematic map can serve as a comprehensive guideline for CA developers for streamlining the existing services as well as developing new systems in accordance with consumer perceptions.

Following the empirical evidence, privacy risk is potentially a major inhibitor of the adoption and use of mental health chatbots. Therefore, CA vendors should provide greater clarity on the nature and use of data in order to reduce subjectively perceived risk of data protection by users. Previous studies suggest that data privacy concerns can be assuaged by providing concise data protection declarations and presenting measures taken for data protection in a layman's language (Featherman & Pavlou 2003). Privacy assurance mechanisms such as certifications by regulatory authorities may further enhance the user's confidence. Further, mandatory access to personal data on the device or user login using social media IDs can also trigger privacy concerns and hence could be avoided. Additionally, it is also recommended they should reduce reliance on third-party messenger platforms owing to the scandals involving data sharing (Weisbaum, 2018), which has reduced consumer trust in these platforms.

Consumer's concern for safety has emerged as a key determinant of use; hence developers should put in place necessary measures to reduce the risk that it poses to the safety and wellbeing of the users. It is evident that currently, the bots have limited systems in place to respond effectively to situations that pose a threat to user safety. Specifically, to counter the risk to life from suicide or self-harming, chatbots should have systems to recognize self-harming intentions and to deal with emergencies. Since the app is used world over, in case of emergency, the helpline numbers relevant to the patient's location should be made available. The apps also should test consumer's perceptions of including information about a trusted friend who could be informed during such emergencies to ensure a more foolproof mechanism (Kretzschmar et al., 2019). Additionally, the overreliance on the mobile may induce addictive behavior and reinforce the tendency to cut off from the real world. Therefore, we recommend that the app should limit the daily usage time beyond unhealthy limits and should have an inbuilt mechanism to suggest and encourage human interactions. Finally, with respect to the users' complaints about chatbots exacerbating their mental health condition, developers should examine the adverse impact of CA-based therapy through clinical research.

As perceived anthropomorphism could potentially influence adoption and usage decisions, the developers should focus on how the design features of CAs can be manipulated to elicit positive outcomes from the consumer. Our research reveals that the personality of CA, Intelligence of CA (conversational responsiveness and comprehension), and perceived empathy are the key design features that could influence adoption decisions. Developers should identify ways to strengthen the positive outcomes (satisfaction, user engagement) by carefully calibrating these key levers. Additionally, designers should be careful that enhancing humanness beyond optimal levels may trigger negative emotional reactions in users, especially in the sensitive context of mental health care.

Additionally, we observe that the CAs have limited cognitive ability and are far from creating the richness of a face to face therapy with the professional. Firstly, the bots were using a limited set of curated scripts for responding to user inputs, making it sound repetitive. Secondly, bots were often restricting the free text user input to avoid conversations spiraling out of control. Even though these strategies help in ensuring user safety, it reduces the conversational responsiveness of the bot. Therefore, the tradeoff between conversational responsiveness and user safety should be further studied to enhance user engagement.

Trust was found to be an influencing factor in the usage of mental health chatbots. We suggest that developers should communicate scientific evidence about the clinical effectiveness of their apps to build initial trust. Initial hype and novelty might get the people to try the app, but ultimately the knowledge-based trust developed during usage experience will determine the adoption and subsequent use (Siau & Wang 2018). Hence building both initial and continuous trust should be a priority of the chatbot developers. Additionally, the trust in providers with respect to data privacy protection should be reinforced through the privacy assurance mechanism, as stated earlier.

The perceived benefit forms the other important determinant of adoption and use. Hence the marketing communications, as well as the service delivery, should prioritize these benefits.

https://aisel.aisnet.org/pajais/vol12/iss2/1
DOI: 10.17705/1pais.12201

With respect to performance expectancy, the inability to handle complex cases was found to limit the perceptions about performance expectancy. In this light, we suggest that the developers should recognize the possibility of users with severe mental disorders seeking services through their application. They should explicitly specify the target audience and the limitations of the app to restrict users with complex mental health issues. Additionally, we suggest that users with complex conditions could be identified and screened using AI-based techniques or some standard screening questionnaires. Further, it was found that apps do not always allow free text input instead provides a list to choose from (fully constrained user input) which restricts the free flow of conversation; developers should work on limiting the use of these 'button press mode of interaction' to enhance the ease of use. Finally, hedonistic value is assumed to be a factor influencing use; the developers can enhance the "humoristic personality" aspect of the bot and communicate this element in their marketing communication.

Furthermore, we believe the four main themes identified in the thematic map have significant implications for the stakeholders like healthcare policymakers and regulatory institutions, who are planning to integrate these systems into formal healthcare delivery systems. We also urge the regulators to look into the concerns raised by users, such as safety risk and privacy risk reported by our study.

We also strongly suggest that the service providers should continuously map and align their expectations on critical determinants for use with those perceived by the consumers to prevent any disappointments. This could help them to identify important product deficiencies or to harmonize communication strategies to mitigate those aspects that may be perceived as critical from the user perspective but are less critical from the viewpoint of the developer.

Limitations and Future Research Directions

The current research has some drawbacks, some of which may represent opportunities for future research. Firstly, the study could be limited by the qualitative approach adopted. We used Netnography followed by thematic analysis of reviews posted by users to arrive at the factors. There are chances that all the factors which have a causal effect on the adoption and use may not have been adequately represented in the reviews. Future researchers may use other methods such as in-depth interviews or focus group discussions with the users (patients) to validate our assertions.

Further, our study did not quantitatively validate the impact of identified factors on the adoption or the interrelationship between the determinants. In the future, quantitative studies can, for example, build upon our theoretical framework to develop a survey to investigate the relative impact of the factors on adoption and to examine the inter factor dynamics. This would make a valuable contribution in generating recommendations for helping the public/private healthcare institutions to address the ambitious goal of improving mental healthcare access in the country. Secondly, we examined user reviews of only two popular mental health CAs; this could be limiting the generalizability of the results. Future studies can include other commercially available CAs to conduct a more elaborate study.

Conclusion

The current study undertook an investigation of factors influencing consumer's adoption and use of automated conversational agents providing mental healthcare services using a qualitative approach based on netnography and thematic analysis. The study developed a comprehensive thematic map, comprising of four main themes perceived risk, perceived benefits, trust, and perceived anthropomorphism and 12 subthemes that visualize the factors that govern the user's adoption and use of mental health chatbots. The insights from the study can guide future research on mental health CA use behavior. It can also serve as a guideline for designers in developing new services or streamlining their existing service offerings to meet customer expectations. It will also aid healthcare policymakers and regulatory institutions, in integrating these mental healthcare CAs into formal healthcare delivery systems.

References

Alalwan, A. A., Dwivedi, Y. K., & Rana, N. P. (2017). Factors influencing adoption of mobile banking by Jordanian bank customers: Extending UTAUT2 with trust. *International Journal of Information Management*, *37*(3), 99-110.

Anand, O., Srivastava, P. R., & Rakshit, A. (2017). Assessment, implication, and analysis of online consumer reviews: a literature review. *Pacific Asia Journal of the Association for Information Systems, 9*(2), 43-74.

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, *85*, 183-189.

Bakker, D., Kazantzis, N., Rickwood, D., & Rickard, N. (2016). Mental health smartphone apps: review and evidence-based recommendations for future developments. *JMIR Mental Health*, *3*(1), e7.

Barak, A., Hen, L., Boniel-Nissim, M., & Shapira, N. A. (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.

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, *1*(1), 71-81.

Bendig, E., Erb, B., Schulze-Thuesing, L., & Baumeister, H. (2019). The next generation: chatbots in clinical psychology and psychotherapy to foster mental health–a scoping review. *Verhaltenstherapie*, 1-13.

Bhugra, D., Tasman, A., Pathare, S., Priebe, S., Smith, S., Torous, J., & First, M. B. (2017). The WPA-lancet psychiatry commission on the future of psychiatry. *The Lancet Psychiatry*, *4*(10), 775-818.

Bird, T., Mansell, W., Wright, J., Gaffney, H., & Tai, S. (2018). Manage your life online: a web-based randomized controlled trial evaluating the effectiveness of a problem-solving intervention in a student sample. *Behavioural and Cognitive Psychotherapy*, *46*(5), 570-582.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101.

Brown, P., Calnan, M., Scrivener, A., & Szmukler, G. (2009). Trust in mental health services: A neglected concept. *Journal of Mental Health*, *18*(5), 449-458.

DOI: 10.17705/1pais.12201

Buxman, K. (1991). Humor in therapy for the mentally ill. *Journal of Psychosocial Nursing and Mental Health Services*, *29*(12), 15-18.

Chung, M., Ko, E., Joung, H., & Kim, S. J. (2018). Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research*.

Connolly, S. L., Miller, C. J., Koenig, C. J., Zamora, K. A., Wright, P. B., Stanley, R. L., & Pyne, J. M. (2018). Veterans' Attitudes Toward Smartphone App Use for Mental Health Care: Qualitative Study of Rurality and Age Differences. *JMIR mHealth and uHealth*, *6*(8), e10748.

D'alfonso, S., Santesteban-Echarri, O., Rice, S., Wadley, G., Lederman, R., Miles, C., ... & Alvarez-Jimenez, M. (2017). Artificial intelligence-assisted online social therapy for youth mental health. *Frontiers in Psychology*, *8*, 796.

Darcy, A. (2017, October 17). Why we need mental health chatbots [Blog post]. Retrieved from https://medium.com/@dralisondarcy/why-we-need-mental-health-chatbots-17559791b2ae

Data miner. (2019). Retrieved from https://data-miner.io/

Davidson, J., & Parr, H. (2008). Virtual trust': online emotional intimacies in mental health support. In Brownlie, J. Greene, A. & Howson, A. (ed.), *Researching trust and health* (45-65). Abingdon, Oxford shire: Routledge.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly,13*(3), 319-340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and intrinsic motivation to use computers in the workplace 1. *Journal of Applied Social Psychology*, *22*(14), 1111-1132.

De Choudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., & Kumar, M. (2016, May). Discovering shifts to suicidal ideation from mental health content in social media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 2098-2110).

Dinev, T., & Hart, P. (2006). Privacy concerns and levels of information exchange: An empirical investigation of intended e-services use. *E-Service*, *4*(3), 25-60.

Donkin, L., Hickie, I. B., Christensen, H., Naismith, S. L., Neal, B., Cockayne, N. L., & 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), e231.

Duarte, P., & Pinho, J. C. (2019). A mixed methods UTAUT2-based approach to assess mobile health adoption. *Journal of Business Research*, *102*, 140-150.

Elson, J. S., Derrick, D., & Ligon, G. (2018). Examining Trust and Reliance in Collaborations between Humans and Automated Agents. In T. Bui (ed.), *Proceedings of the 51st Hawaii International Conference on System Sciences*, Hawaii International Conference on System Sciences.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864.

Epley, N., Waytz, A., Akalis, S., & Cacioppo, J. T. (2008). When we need a human: Motivational determinants of anthropomorphism. *Social Cognition*, *26*(2), 143-155.

Facebook (2019). *Facebook data policy*. Retrieved from https://www.facebook.com/full_data_use_policy

Featherman, M. S., & Pavlou, P. A. (2003). Predicting e-services adoption: a perceived risk facets perspective. *International Journal of Human-Computer Studies*, *59*(4), 451-474.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, *5*(1), 80-92.

Firth, J., & Torous, J. (2015). Smartphone apps for schizophrenia: a systematic review. *JMIR mHealth and uHealth*, *3*(4), e102.

Fitzpatrick, K. 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), e19.

Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2018). Using Psychological Artificial Intelligence (Tess) to Relieve Symptoms of Depression and Anxiety: Randomized Controlled Trial. *JMIR Mental Health*, *5*(4), e64.

Gaffney, H., Mansell, W., & Tai, S. (2019). Conversational Agents in the Treatment of Mental Health Problems: Mixed-Method Systematic Review. *JMIR Mental Health*, *6*(10), e14166.

Gilbody, S., Littlewood, E., Hewitt, C., Brierley, G., Tharmanathan, P., Araya, R., ... & Kessler, D. (2015). Computerised cognitive behaviour therapy (cCBT) as treatment for depression in primary care (REEACT trial): large scale pragmatic randomised controlled trial. *BMJ*, *351*, h5627.

Ginn, S., & Horder, J. (2012). “One in four” with a mental health problem: the anatomy of a statistic. *BMJ*, *344*, e1302.

Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, *97*, 304-316.

Goulding, C. 2000. Grounded theory methodology and consumer behavior, procedures, practice and pitfalls. *Advances in Consumer Research, 27*(1), 261-266.

Gulliver, A., Griffiths, K. M., & Christensen, H. (2010). Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. *BMC Psychiatry*, *10*(1), 113.

Gursoy, D., Chi, O. H., Lu, L., & Nunkoo, R. (2019). Consumers acceptance of artificially intelligent (AI) device use in service delivery. *International Journal of Information Management*, *49*, 157-169.

Han, S. & Yang, H. (2018). Understanding adoption of intelligent personal assistants: A parasocial relationship perspective. *Industrial Management & Data Systems*, *118*(3), 618–636.

Hengstler, M., Enkel, E., & Duelli, S. (2016). Applied artificial intelligence and trust—The case of autonomous vehicles and medical assistance devices. *Technological Forecasting and Social Change*, *105*, 105-120.

Hill, J., Ford, W. R., & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human– human online conversations and human–chatbot conversations. *Computers in Human Behavior, 49*, 245–250.

Ho, C. C., & MacDorman, K. F. (2017). Measuring the uncanny valley effect. *International Journal of Social Robotics*, *9*(1), 129-139.

Hollis, C., Sampson, S., Simons, L., Davies, E. B., Churchill, R., Betton, V., ... & Kabir, T. (2018). Identifying research priorities for digital technology in mental health care: results of the James Lind Alliance Priority Setting Partnership. *The Lancet Psychiatry*, *5*(10), 845-854.

https://aisel.aisnet.org/pajais/vol12/iss2/1
DOI: 10.17705/1pais.12201

Hoque, R., & Sorwar, G. (2017). Understanding factors influencing the adoption of mHealth by the elderly: An extension of the UTAUT model. *International Journal of Medical Informatics*, *101*, 75-84.

Hruschka, D. J., Schwartz, D., St. John, D. C., Picone-Decaro, E., Jenkins, R. A., & Carey, J. W. (2004). Reliability in coding open-ended data: Lessons learned from HIV behavioral research. *Field Methods*, *16*(3), 307-331.

Huang, H. Y., & Bashir, M. (2017). Users' adoption of mental health apps: examining the impact of information cues. *JMIR mHealth and uHealth*, *5*(6), e83.

Huang, J., Li, Q., Xue, Y., Cheng, T., Xu, S., Jia, J., & Feng, L. (2015, May). Teenchat: a chatterbot system for sensing and releasing adolescents' stress. In *International Conference on Health Information Science* (pp. 133-145). Springer, Cham.

Kari, T., Koivunen, S., Frank, L., Makkonen, M., & Moilanen, P. (2016). Critical experiences during the implementation of a self-tracking technology. In *Proceedings of the 20th Pacific Asia Conference on Information Systems 2016.* Chiyayi, Taiwan: AIS.

Kazdin, A. E., & Rabbitt, S. M. (2013). Novel models for delivering mental health services and reducing the burdens of mental illness. *Clinical Psychological Science*, *1*(2), 170-191.

Kim, D., Park, K., Park, Y., & Ahn, J. H. (2019). Willingness to provide personal information: Perspective of privacy calculus in IoT services. *Computers in Human Behavior*, *92*, 273-281.

Kozinets, R. V. (2002). The field behind the screen: Using netnography for marketing research in online communities. *Journal of Marketing Research*, *39*(1), 61-72.

Kretzschmar, K., Tyroll, H., Pavarini, G., Manzini, A., Singh, I., & NeurOx Young People's Advisory Group. (2019). Can Your Phone Be Your Therapist? Young People's Ethical Perspectives on the Use of Fully Automated Conversational Agents (Chatbots) in Mental Health Support. *Biomedical Informatics Insights*, *11*, 1178222619829083.

Lankton, N. K., McKnight, D. H., & Tripp, J. (2015). Technology, humanness, and trust: Rethinking trust in technology. *Journal of the Association for Information Systems*, *16*(10), 880-918.

Li, H., Wu, J., Gao, Y., & Shi, Y. (2016). Examining individuals' adoption of healthcare wearable devices: An empirical study from privacy calculus perspective. *International Journal of Medical Informatics*, *88*, 8-17.

Li, X., Hess, T. J., & Valacich, J. S. (2008). Why do we trust new technology? A study of initial trust formation with organizational information systems. *The Journal of Strategic Information Systems*, *17*(1), 39-71.

Lipschitz, J., Miller, C. J., Hogan, T. P., Burdick, K. E., Lippin-Foster, R., Simon, S. R., & Burgess, J. (2019). Adoption of Mobile Apps for Depression and Anxiety: Cross-Sectional Survey Study on Patient Interest and Barriers to Engagement. *JMIR Mental Health*, *6*(1), e11334.

Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., ... & AlMazroa, M. A. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, *380*(9859), 2095-2128.

Lucas, G. M., Gratch, J., King, A., & Morency, L. P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, *37*, 94-100.

Luo, X., Li, H., Zhang, J., & Shim, J. P. (2011). Examining multi-dimensional trust and multi-faceted risk in initial acceptance of emerging technologies: An empirical study of mobile banking services. *Decision Support Systems*, *49*(2), 222-234.

MacDorman, K. F., & Chattopadhyay, D. (2017). Categorization-based stranger avoidance does not explain the uncanny valley effect. *Cognition*, *161*, 132–135.

Madan, K. & Yadav, R. (2018). Understanding and predicting antecedents of mobile shopping adoption: A developing country perspective. *Asia Pacific Journal of Marketing and Logistics*, *30*(1), 139–162.

Martinez-Martin, N., & Kreitmair, K. (2018). Ethical issues for direct-to-consumer digital psychotherapy apps: addressing accountability, data protection, and consent. *JMIR Mental Health*, *5*(2), e32.

Matt, C., Becker, M., Kolbeck, A., & Hess, T. (2019). Continuously Healthy, Continuously Used?–A Thematic Analysis of User Perceptions on Consumer Health Wearables. *Pacific Asia Journal of the Association for Information Systems*, *11*(1), 108-132.

McCarthy, J. (2007). From here to human-level AI. *Artificial Intelligence*, *171*(18), 1174-1182.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, *22*(3), 276-282.

Mcknight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on Management Information Systems*, *2*(2), 1-25.

Melián-González, S., Gutiérrez-Taño, D., & Bulchand-Gidumal, J. (2019). Predicting the intentions to use chatbots for travel and tourism. *Current Issues in Tourism*, 1-19.

Merry, S. N., Stasiak, K., Shepherd, M., Frampton, C., Fleming, T., & Lucassen, M. F. (2012). The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. *BMJ*, *344*, e2598.

Microsoft Corporation. (2018). *The future computed: Artificial Intelligence and its role in society*. Retrieved from https://news.microsoft.com/uploads/2018/02/The-Future-Computed_2.8.18.pdf

Miner, A. S., Milstein, A., & Hancock, J. T. (2017). Talking to machines about personal mental health problems. *JAMA*, *318*(13), 1217-1218.

Miner, A. S., Shah, N., Bullock, K. D., Arnow, B. A., Bailenson, J., & Hancock, J. (2019). Key Considerations for Incorporating Conversational AI in Psychotherapy. *Frontiers in Psychiatry*, *10*, 746.

Mohr, D. C., Weingardt, K. R., Reddy, M., & Schueller, S. M. (2017). Three problems with current digital mental health research... and three things we can do about them. *Psychiatric Services*, *68*(5), 427-429.

Molteni, M. (2017, June 08). The chatbot therapist will see you now. Wired. Retrieved from https://www.wired.com/2017/06/facebook-messenger-woebot-chatbot-therapist

Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research*, *2*(3), 192-222.

Mora-Ripoll, R. (2010). The therapeutic value of laughter in medicine. *Alternative Therapies in Health & Medicine*, *16*(6), 56-64.

Morris, R. R., Kouddous, K., Kshirsagar, R., & Schueller, S. M. (2018). Towards an artificially empathic conversational agent for mental health applications: System design and user perceptions. *Journal of Medical Internet Research*, *20*(6), e10148.

https://aisel.aisnet.org/pajais/vol12/iss2/1
DOI: 10.17705/1pais.12201

Nambisan, P. (2011). Information seeking and social support in online health communities: impact on patients' perceived empathy. *Journal of the American Medical Informatics Association*, *18*(3), 298-304.

Naslund, J. A., Aschbrenner, K. A., Araya, R., Marsch, L. A., Unützer, J., Patel, V., & Bartels, S. J. (2017). Digital technology for treating and preventing mental disorders in low-income and middle-income countries: a narrative review of the literature. *The Lancet Psychiatry*, *4*(6), 486-500.

Nienhuis, J. B., Owen, J., Valentine, J. C., Winkeljohn Black, S., Halford, T. C., Parazak, S. E., ... & Hilsenroth, M. (2018). Therapeutic alliance, empathy, and genuineness in individual adult psychotherapy: A meta-analytic review. *Psychotherapy Research*, *28*(4), 593-605.

Novak, J. D., & Cañas, A. J. (2006). The theory underlying concept maps and how to construct them. *Florida Institute for Human and Machine Cognition*, *1*(1), 1-31.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods*, *16*(1), 1–13.

Paiva, A., Leite, I., Boukricha, H., & Wachsmuth, I. (2017). Empathy in virtual agents and robots: a survey. *ACM Transactions on Interactive Intelligent Systems*, *7*(3), 1-40.

Payne, E. M., Peltier, J. W., & Barger, V. A. (2018). Mobile banking and AI-enabled mobile banking: The differential effects of technological and non-technological factors on digital natives' perceptions and behavior. *Journal of Research in Interactive Marketing*, *12*(3), 328–346.

Ravoniarison, A., & Benito, C. (2019). Mobile games: players' experiences with in-app purchases. *Journal of Research in Interactive Marketing*, *13*(1), 62–78.

Saffarizadeh, K., Boodraj, M., & Alashoor, T. M. (2017). Conversational Assistants: Investigating Privacy Concerns, Trust, and Self-Disclosure. In *Proceedings of the 2017 International Conference on Information Systems*. Seoul, Korea.

Saldana, J. (2015). *The coding manual for qualitative researchers*. Thousand Oaks, California: Sage.

Sarikaya, R. (2017). The technology behind personal digital assistants: An overview of the system architecture and key components. *IEEE Signal Processing Magazine*, *34*(1), 67-81.

Schroeder, J., & Schroeder, M. (2018). Trusting in machines: How mode of interaction affects willingness to share personal information with machines. In *Proceedings of the 51st Hawaii International Conference on System Sciences* (pp. 472-480). Hawaii, USA.

Schueller, S. M., Neary, M., O'Loughlin, K., & Adkins, E. C. (2018). Discovery of and interest in health apps among those with mental health needs: survey and focus group study. *Journal of Medical Internet Research*, *20*(6), e10141.

Schuetzler, R. M., Grimes, G. M., Giboney, J. S., & Nunamaker Jr, J. F. (2018). The influence of conversational agents on socially desirable responding. In *Proceedings of the 51st Hawaii International Conference on System Sciences* (pp. 283-292). Hawaii, USA.

Seeger, A. M., Pfeiffer, J., & Heinzl, A. (2018). Designing Anthropomorphic Conversational Agents: Development and Empirical Evaluation of a Design Framework. In T. Bui (Ed.), *Proceedings of the 39th International Conference on Information Systems*, San Francisco, USA.

Shawar, A., & Atwell, E. S. (2007). Chatbots: are they really useful? *Journal for Language Technology and Computational Linguistics*, *22*(1), 29-49.

Sheehan, B. T. (2018). Customer service chatbots: Anthropomorphism, adoption and word of mouth (Doctoral dissertation). Queensland University of Technology. Retrieved from https://eprints.qut.edu.au/121188/

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, *22*(2), 63-75.

Shum, H. Y., He, X. D., & Li, D. (2018). From Eliza to XiaoIce: challenges and opportunities with social chatbots. *Frontiers of Information Technology & Electronic Engineering*, *19*(1), 10-26.

Siau, K., & Wang, W. (2018). Building trust in artificial intelligence, machine learning, and robotics. *Cutter Business Technology Journal*, *31*(2), 47-53.

Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: an interdisciplinary review. *MIS Quarterly*, *35*(4), 989-1016.

Söllner, M., Hoffmann, A., & Leimeister, J. M. (2016). Why different trust relationships matter for information systems users. *European Journal of Information Systems*, *25*(3), 274-287.

Stawarz, K., Preist, C., Tallon, D., Wiles, N., & Coyle, D. (2018). User experience of cognitive behavioral therapy apps for depression: an analysis of app functionality and user reviews. *Journal of Medical Internet Research*, *20*(6), e10120.

Ta, V., Griffith, C., Boatfield, C., Wang, X., Civitello, M., Bader, H., … Loggarakis, A. (2020). User Experiences of Social Support from Companion Chatbots in Everyday Contexts: Thematic Analysis. *Journal of Medical Internet Research*, *22*(3), e16235.

Tarhini, A., Al-Busaidi, K. A., Mohammed, A. B., & Maqableh, M. (2017). Factors influencing students' adoption of e-learning: a structural equation modeling approach. *Journal of International Education in Business*, *10*(2), 164–182.

Tegmark, M. (2017). *Life 3.0: Being Human in the Age of Artificial Intelligence*. New York, NY: Alfred A. Knopf.

Torous, J., Friedman, R., & Keshavan, M. (2014). Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. *JMIR mHealth and uHealth*, *2*(1), e2.

Trivedi, J. (2019). Examining the customer experience of using banking Chatbots and its impact on brand love: the moderating role of perceived risk. *Journal of Internet Commerce*, *18*(1), 91-111.

Van Doorn, J., Mende, M., Noble, S. M., Hulland, J., Ostrom, A. L., Grewal, D., & Petersen, J. A. (2017). Domo arigato Mr. Roboto: Emergence of automated social presence in organizational frontlines and customers' service experiences. *Journal of Service Research*, *20*(1), 43-58.

Vasa, R., Hoon, L., Mouzakis, K., & Noguchi, A. (2012, November). A preliminary analysis of mobile app user reviews. In *Proceedings of the 24th Australian Computer-Human Interaction Conference* (pp. 241-244). Melbourne, Australia.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425-478.

Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly, 36*(1), 157-178.

Venkatesh, V., Thong, J. Y., & Xu, X. (2016). Unified theory of acceptance and use of technology: A synthesis and the road ahead. *Journal of the Association for Information Systems*, *17*(5), 328-376.

https://aisel.aisnet.org/pajais/vol12/iss2/1
DOI: 10.17705/1pais.12201

Vigo, D., Thornicroft, G., & Atun, R. (2016). Estimating the true global burden of mental illness. *The Lancet Psychiatry*, *3*(2), 171-178.

Wang, W. (2017). Smartphones as Social Actors? Social dispositional factors in assessing anthropomorphism. *Computers in Human Behavior*, *68*, 334-344.

Wang, X., White, L., Chen, X., Gao, Y., Li, H., & Luo, Y. (2015). An empirical study of wearable technology acceptance in healthcare. *Industrial Management & Data Systems*, *115*(9), 1704–1723.

Wang, Y. S., Lee, W. L., & Hsu, T. H. (2017). Using netnography for the study of role-playing in female online games: Interpretation of situational context model. *Internet Research*, *27*(4), 905–923.

Waytz, A., Cacioppo, J., & Epley, N. (2010). Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science*, *5*(3), 219-232.

Weisbaum, H. (2018, April 19). Trust in Facebook Has Dropped by 66 Percent Since the Cambridge Analytica Scandal. *NBC News*. Retrieved from https://www.nbcnews.com/business/consumer/trust-facebook-has-dropped-51-percent-cambridge-analytica-scandal-n867011

Whiteford, H. A., Ferrari, A. J., Degenhardt, L., Feigin, V., & Vos, T. (2015). The global burden of mental, neurological and substance use disorders: an analysis from the Global Burden of Disease Study 2010. *PloS One*, *10*(2), e0116820.

WHO (2017). *Depression and other common mental disorders global health estimates*. Retrieved from http://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf

WHO (2018). *Mental health atlas 2017*. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/272735/9789241514019-eng.pdf?ua=1

Woebot (2019). Woebot website. Retrieved from https://woebot.io/

Wysa (2019). Wysa website. Retrieved from https://www.wysa.io/

Yang, C. C., Tang, X., Wong, Y. C., & Wei, C. P. (2010). Understanding online consumer review opinions with sentiment analysis using machine learning. *Pacific Asia Journal of the Association for Information Systems*, *2*(3), 73-89.

Yang, Z., & Fang, X. (2004). Online service quality dimensions and their relationships with satisfaction: A content analysis of customer reviews of securities brokerage services. *International Journal of Service Industry Management*, *15*(3), 302–326.

Yuan, S., Ma, W., Kanthawala, S., & Peng, W. (2015). Keep using my health apps: Discover users' perception of health and fitness apps with the UTAUT2 model. *Telemedicine and e-Health*, *21*(9), 735-741.

Zaidan, S., & Roehrer, E. (2016). Popular Mobile Phone Apps for Diet and Weight Loss: A Content Analysis. *JMIR mHealth and uHealth*, *4*(3), e80.

Zarouali, B., Van den Broeck, E., Walrave, M., & Poels, K. (2018). Predicting consumer responses to a chatbot on Facebook. *Cyberpsychology, Behavior, and Social Networking*, *21*(8), 491-497.

About the Authors

Ashish Viswanath Prakash is a Doctoral candidate and a Senior research fellow at the Vinod Gupta School of Management, Indian Institute of Technology, Kharagpur, India. His research interest lies in digital innovation, IT in healthcare, Technology adoption, and Human-Computer Interaction. His Ph.D. project focuses on the adoption of Artificial Intelligence in healthcare from the perspectives of technology, policy, and management. He has published in the Journal of International Education in Business.

Dr. Saini Das is an Assistant Professor at the Vinod Gupta School of Management, Indian Institute of Technology, Kharagpur, India. Her major research interests are in managing information security risks in networks, management information systems (MIS), e-commerce technology and applications, data privacy, digital piracy, data analytics and artificial intelligence (AI). She has authored publications in several international journals of repute, including Decision Support Systems, Behaviour & Information Technology, Information Systems Frontiers, Journal of Global Information Technology Management, and Journal of Information Privacy and Security.