Insights and lessons learned from trialling a mental health chatbot in the wild

Published in:
26th IEEE Symposium on Computers and Communications, ISCC 2021

Publication Status:
Published (in print/issue): 15/12/2021

DOI:
10.1109/ISCC53001.2021.9631395

Document Version
Author Accepted version

Insights and lessons learned from trialling a mental health chatbot in the wild

Courtney Potts
School of Computing
Ulster University
Newtownabbey, UK
c.potts@ulster.ac.uk

Raymond Bond
School of Computing
Ulster University
Newtownabbey, UK
rb.bond@ulster.ac.uk

Maurice D. Mulvenna
School of Computing
Ulster University
Newtownabbey, UK
md.mulvenna@ulster.ac.uk

Edel Ennis
School of Psychology
Ulster University
Derry-Londonderry, UK
e.ennis@ulster.ac.uk

Andrea Bickerdike
Department of Sport, Leisure and Childhood Studies
Munster Technological University
Cork, Ireland
andrea.bickerdike@cit.ie

Edward K. Coughlan
Department of Sport, Leisure and Childhood Studies
Munster Technological University)
Cork, Ireland
edward.coughlan@cit.ie

Thomas Broderick
Department of Sport, Leisure and Childhood Studies
Munster Technological University)
Cork, Ireland
thomas.broderick@cit.ie

Con Burns
Department of Sport, Leisure and Childhood Studies
Munster Technological University)
Cork, Ireland
con.burns@cit.ie

Michael F. McTear
School of Computing
Ulster University
Newtownabbey, UK
mf.mctear@ulster.ac.uk

Lauri Kuosmanen
Department of Nursing Science
University of Eastern Finland
Kuopio, Finland
lauri.kuosmanen@uef.fi

Heidi Nieminen
Department of Nursing Science
University of Eastern Finland
Kuopio, Finland
heidi.nieminen@uef.fi

Kyle A. Boyd
School of Art
Ulster University
Belfast, UK
ka.boyd@ulster.ac.uk

Brian Cahill
Nimbus Research Centre
Munster Technological University
Cork, Ireland
brian.cahill@cit.ie

Alex Vakaloudis
Nimbus Research Centre
Munster Technological University
Cork, Ireland
alex.vakaloudis@cit.ie

Indika Dhanapala
Nimbus Research Centre
Munster Technological University
Cork, Ireland
indika.dhanapala@cit.ie

Anna-Kaisa Vartiainen
Department of Nursing Science
University of Eastern Finland
Kuopio, Finland
anna-kaisa.vartiainen@uef.fi

Catrine Kostenius
Department of Health Sciences
Luleå University of Technology
Lulea, Sweden
catrine.kostenius@ltu.se

Martin Malcolm
Health Intelligence & Information Services
NHS Western Isles
Stornoway, UK
martin.malcolm2@nhs.scot

***Abstract*— This study reports on the development and in the wild trialling of a chatbot (ChatPal) which promotes good mental wellbeing. A stakeholder-centered approach for design was adopted where end users, mental health professionals and service users were involved in the design which was centered around positive psychology. In the wild usage of the chatbot was explored from Jul-20–Mar-21. Exploratory analyses of usage metrics were carried out using the event log data. User tenure, unique usage days, total chatbot interactions and average daily interactions were used in K-means clustering to identify user archetypes. The chatbot was used by a variety of age groups (18-65+) and genders, mainly those living in Ireland. K-means clustering identified three clusters: sporadic users (n=4), frequent transient users (n=38) and abandoning users (n=169) each with distinct usage characteristics. This study highlights the importance of event log data analysis for making improvements to the mental health chatbot.**

Keywords*— *Conversational user interfaces, event log, eHealth, mental wellbeing, co-design, COVID-19

I. Introduction

Chatbots or conversational user interfaces are popular eHealth tools that have the potential to promote good mental health in the general population. These digital wellbeing interventions have shown promise in terms of efficacy [1]–[3], availability and accessibility [4], [5]. Mental health chatbots may be geared towards a variety of outcomes including psychoeducation, self-management of mental health and wellbeing by tracking mood or monitoring symptom change, and promoting help seeking behaviour [1]. In addition to supporting those with mental ill health, digital technologies such as chatbots are also considered to have potential for preventing mental health problems and for improving the overall mental health of the population [6]. Recent advances in Artificial Intelligence (AI) and Natural Language

Processing (NLP) have enabled a rapid increase in the number of chatbots developed in this area. However this has also led to concerns that studies have typically not employed Randomised Control Trial (RCT) methodology, included a variety of sample groups or small sample sizes, have high dropout or attrition rates, used various outcome measures and inconsistent follow-up protocols, and used various formats of intervention and comparison conditions [1]. However, Calvo and colleagues note that an RCT may not be the most appropriate methodology as the strength of digital technology may lie in the ability to provide an individual/personalised intervention or approach [6]. Few studies have explored the use of mental health chatbots among a non-clinical population [5], [7], [8]. However, this may be difficult as research into the benefits of these chatbots is unlikely to yield a statistically significant effect in those who already have good mental health. Nonetheless, chatbots could play an important role in building and maintaining good mental health across the population, especially during the ongoing COVID-19 pandemic [9], [10]. Further research is necessary to determine how digital technology can be best used in the mental health sector and what developments or limitations need to be incorporated to make the intervention acceptable, effective, safe and financially viable [1], [11].

The present study is part of a larger project known as ChatPal (https://chatpal.interreg-npa.eu/). The remit of the broader 'ChatPal' project extends to the development, trialling, roll-out and evaluation of a chatbot that promotes positive mental wellbeing of individuals, designed for use in Northern Periphery Areas (NPAs) (Northern Ireland, Ireland, Scotland, Finland and Sweden). The project has three main objectives: to understand the digital mental health requirements of citizens in rural areas; to co-create and pilot a multilingual chatbot that is effective for providing a blended digital mental health service across NPAs; and to inform and increase awareness and attitudes of mental healthcare professionals regarding the use of eHealth tools (chatbots) to augment and improve mental health service provision.

This study reports on the initial prototype which was developed and released early to support people at the beginning of the COVID-19 pandemic. The content of the prototype was based around positive psychology, containing elements of psychological well-being and happiness, known as the PERMAH model which includes Positive emotions, Engagement, Relationships, Meaning, Accomplishment and Health (PERMAH) [12]–[14]. To date, there is limited evidence pertaining to chatbots that have used content from positive psychology [5], [15].

The aim of this research is to report on the engagement with the ChatPal chatbot during the pandemic and subsequent insights gained from an 'in the wild' trial. Thus, this study seeks to answer four research questions: What were the characteristics of individuals that used the ChatPal chatbot? How did people interact with the ChatPal chatbot during the pandemic? What user groups can be identified using unsupervised machine learning (clustering)? What lessons can be learned from trialling a mental health chatbot?

II. Methods

This study has obtained ethical approval from the School of Psychology Filter (Ethics) Committee, Ulster University.

A. Chatbot Co-design Methodology

A stakeholder-centered approach for responsible mental health chatbot design was adopted. At the beginning and throughout the ChatPal project, mental health professionals were surveyed to ascertain attitudes towards prescribing eHealth tools and chatbots [16]. Needs analysis workshops took place in tandem across Northern Ireland, Republic of Ireland, Scotland, Sweden and Finland to gather user requirements. These workshops were attended by the general population (n=40), mental health professionals (n=27) and mental health service users (n=11) to gain a broad picture on what people need and/or want to see in the chatbot in terms of content and personality. Content was developed based on the PERMAH positive psychology model [17] to address user requirements. Multidisciplinary teams worked together to design and create dialogue flow charts based on the PERMAH model which were then implemented into the chatbot.

ChatPal can maintain a basic dialogue with users around how to maintain positive emotions and good physical health, develop relationships, as well as engendering a sense of accomplishment and meaning. The majority of dialogues within ChatPal are scripted conversations with predefined responses for the user to select. A free text input option is available for users to express what they felt grateful for (gratitude statements). Users can select to complete the World Health Organisation Well Being Index (WHO-5) [18] which is a short assessment of overall perceived sense of wellbeing. The WHO-5 is scored out of 25, with raw scores multiplied by 4 to give a total out of 100. Higher scores relate to better mental wellbeing, and a score of 52 or below is indicative of poor wellbeing. At any time while using the chatbot, users could select 'need to go' which would indicate the end of the conversation.

B. Chatbot Architecture

The implementation of the ChatPal chatbot is based on Rasa and PhoneGap frameworks. An overview of the design can be seen in Fig. 1. Rasa is an AI assisted framework for building contextual chatbots and provides infrastructure and tools necessary for high-performing, resilient, proprietary contextual assistants. PhoneGap is an open-source framework for developing cross-platform mobile applications. The backend of ChatPal uses the Rasa framework. Upon receiving user inputs from the ChatPal app, the Rasa stack forwards the user input to the Rasa Natural Language Understanding (NLU) unit, which extracts user intentions and relevant metadata out of the input. These intentions and metadata are called intents and entities within the Rasa framework. Once the intents and entities are identified, corresponding responses need to be carried out which are determined by the Rasa core. Rasa supports various responses including text messages, buttons, images and reminders. Rasa also allows custom actions which are executed separately in an action server which is run with the help of the Rasa SDK. The SQLAlchemy python library, which uses Postgres SQL queries, was used as the custom actions require database access. The PhoneGap framework is used as the front end of ChatPal. The communication between the ChatPal backend is done with HTTP requests/ responses. The ChatPal chatbot prototype was released in English on the Google Play store on 30th June 2020. A web version of the chatbot was also made available at this time. Participants were not recruited to use the app, instead it was shared on social media and various online news articles were published with links to download the chatbot.

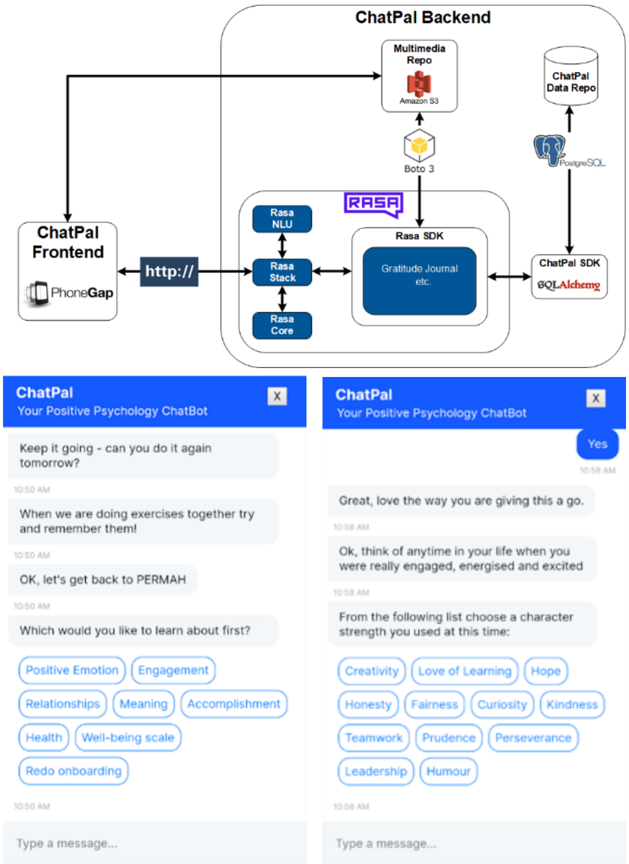

Fig. 1. ChatPal architecture overview, screenshot of menu and example dialogue from "Engagement" conversation.

C. *Real World Testing and Data Analysis*

At onboarding users consented to the storage and analysis of anonymous chatbot data. Event logging was built into the chatbot so that every interaction was anonymously recorded along with the date-time stamp and a unique user ID. The user ID was automatically assigned to each user. However, the assignment of user IDs did not work correctly at the beginning of the study as users were given a new ID every time they opened the app. This issue was corrected on 7 October 2020. Event log data were analysed from 1 July 2020 - 1 March 2021 (8 months, 244 days). Within the event logs, the beginning of conversations were automatically labelled as the start of a session. To indicate the end of a session the user had to select 'need to go' and this was recorded in the event logs. Following this, an exit message would be displayed to the user which contained a link to an external survey with the Chatbot Usability Questionnaire (CUQ) [19] to measure feedback. The methodology followed the Health Interaction Log Data Analytics (HILDA) pipeline [20]. All analyses were carried out using R Studio, R version 3.6.0. R libraries were used for the analysis including; dplyr for data wrangling and ggplot2 for data visualisation. Overall, monthly, and daily chatbot interactions were explored over the duration of the 8 month period along with usage over the days of the week and hours of the day. Gender, age and country of origin were computed for the users that chose to provide this information. Engagement with features was measured by the number of interactions within each of the PERMAH topics and summary statistics were computed for the WHO-5 scale. Free text from user input gratitude statements was visualised using a word cloud. For each individual user, chatbot usage from 7th October 2020 - 1st March 2021 was explored. A new dataset was created from the event logs which included the user tenure (duration of use in days), number of unique days, total interactions, and average daily interactions for each user. These features were used in k-means clustering, an unsupervised machine learning technique, which groups a set of data points together in such a way that data in one group has similar properties and is well-defined from other groups. In this study K-means clustering was used to characterise chatbot users into distinct groups. The R package NbClust was used to determine the number of clusters by comparing 30 indices to select the number of clusters according to the majority rule.

The ChatPal consortium acknowledges the support provided by the Interreg VB Northern Periphery & Arctic Programme under the grant for Conversational Interfaces Supporting Mental Health and Wellbeing of People in Sparsely Populated Areas (ChatPal) project number 345.

III. Results

A. *Interactions Over Time*

The total number of interactions for all users over the 8-month period was 25,713 (mean = 106 daily interactions, SD = 171; mean = 3,200 monthly interactions, SD = 1,430). There was an overall strong negative correlation in monthly interactions over time ($r = -0.82$, $p = 0.012$). Weekdays were the most popular time users interacted with the chatbot, with 20% of all interactions on a Thursday. Generally, the fewest interactions were recorded on Saturdays (n = 1,669, 6%) and Sundays (n = 2,175, 8%). Across hours of the day, the most common time to interact with the chatbot was 11am (n = 3,133, 12%).

B. *User Demographics*

Across the 8-month data collection period, a subset of users provided their gender (n = 103), age (n = 113) and country of origin (n = 104) as outlined in Table 1. From the information provided by users, a higher proportion of males downloaded and used the chatbot compared to females. Those who downloaded the chatbot and selected that they were under 18 were not permitted to continue using the chatbot (n = 7). Users' ages varied from 18 to over 65, with the largest group being those in the 25-34 age range. Roughly a third of users were from Northern Ireland, a quarter were from Ireland while the rest were from Scotland, Sweden, Finland or other regions (Table 1).

TABLE I. Known User Demographics

Gender, n (%)	**Male**	58 (56%)
	Female	40 (39%)
	Transgender (female to male)	1 (1%)
	Other/ prefer not to say	4 (4%)
Age, n (%)	**<18**	7 (6%)
	18-24	20 (18%)
	25-34	30 (27%)
	35-44	22 (20%)
	45-54	15 (14%)
	55-64	10 (9%)
	>65	7 (6%)
Country, n (%)	**Northern Ireland**	33 (32%)
	Ireland	26 (25%)
	Scotland	2 (2%)
	Finland	6 (6%)
	Sweden	6 (6%)
	Other/ Prefer not to say	31 (29%)

C. Chatbot Features

The chatbot offered six main topics of conversation with the user under the 'PERMAH' model which included Positive Emotion, Engagement, Relationships, Meaning, Accomplishment, and Health. The frequency of interaction with each of these topics within the chatbot was explored (Fig. 2). Positive Emotion was the most commonly selected PERMAH dialogue, followed by Meaning, Health, Engagement and Relationships while Accomplishment was only picked once (Fig. 2). A total of 25 individuals completed the WHO-5 scale, with scores ranging from 12 to 84, (mean = 50.4, SD = 21.0). Mean score was below threshold of 52 indicating poor mental wellbeing on average.

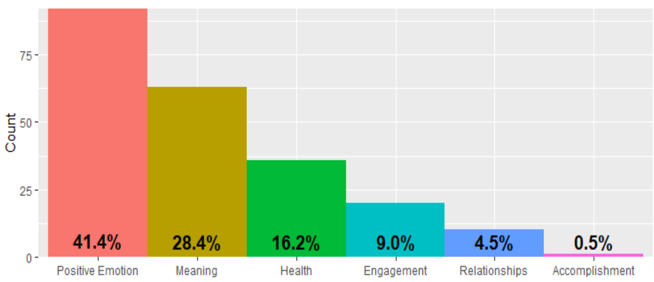

Fig. 2. Frequency of interactions with each of the 'PERMAH' dialogues within the chatbot.

As part of the Positive Emotion dialogue, users could input gratitude statements. Enjoying one's job, and having great family and friends were the most expressed reasons to be grateful.

The beginning of a session was recorded 496 times (1.9% of all interactions), however the end of a session was recorded only 69 times, (0.2% interactions). At the end of a session, the exit message which contained a survey link to provide feedback on ChatPal was displayed, but only 7 people fully completed the feedback questionnaire.

D. Chatbot Usage and Clustering

The subsequent analyses were carried out over a 4-month period from October 2020 to March 2021. During this time there were a total of 10,311 chatbot interactions logged from 211 unique users. Characteristics of behaviour were used as features in K-means clustering to identify different groups of chatbot users. The features were as follows: *User tenure*: duration of chatbot use in terms of time in days between first and last chatbot interaction; *Unique days*: Total number of unique days the person used the chatbot; *Total interactions*: Total number of chatbot interactions over duration of use; *Average interactions (day)*: mean number of daily chatbot interactions. Based on the metrics tested and according to the majority rule, 8 indices proposed 3 as the best number of clusters. Therefore, users were segregated into 3 distinct groups (Fig. 4).

Cluster 1 "sporadic users" were a small group of individuals (n = 4), who used the chatbot on average for 7 unique days over a three month period, logging a high number of interactions during this time (Table 2, Fig. 4). Cluster 2 "abandoning users" was the largest cluster made up of the majority of chatbot users (n = 169). These people had the fewest number of interactions and shortest tenure, on average only using the chatbot for 1 day (Table 2, Fig. 4). Chatbot users in cluster 3 (n = 38), termed "Frequent transient users", typically interacted extensively with the chatbot over two days in a 9-day period, although user tenure varied greatly. There were no distinct patterns in gender, age, and country across clusters 1, 2 and 3 given these details were largely unknown for most users (Table 2, Fig. 4).

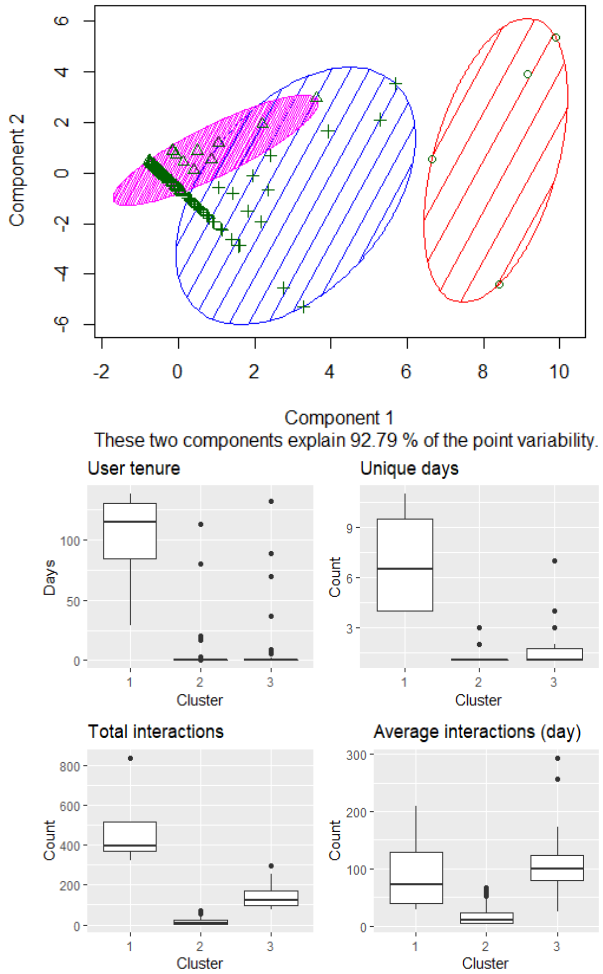

Fig. 3. Visualisation of three clusters shown in two principal components and boxplots of each feature per cluster.

TABLE II. USER CHARACTERISTICS FOR EACH CLUSTER

Cluster		1: "Sporadic users"	2: "Abandoning users"	3: "Frequent transient users"
Size		4	169	38
User tenure (days)	Mean (SD)	99.0 (49.1)	1.5 (10.9)	9.3 (27.7)
Unique days	Mean (SD)	7 (3.5)	1.1 (0.3)	1.6 (1.2)
Total interactions	Mean (SD)	488 (236)	18 (17.4)	140 (54.8)
Average daily interactions	Mean (SD)	96.0 (82.1)	16.6 (15.4)	110 (52.8)
Gender, n (%)	Female Male Trans (female to male) Unknown	1 (25%) - - 3 (75%)	4 (2%) 3 (2%) - 162 (96%)	9 (24%) 16 (42%) 1 (3%) 12 (31%)
Age, n (%)	18-24 25-34 35-44 45-54 55-64 >65 Unknown	2 (50%) 1 (25%) - - 1 (25%) - -	6 (4%) 1 (<1%) 4 (2%) 2 (1%) 1 (<1%) - 155 (92%)	7 (18%) 11 (29%) 5 (13%) 4 (11%) - 1 (3%) 10 (26%)
Country, n (%)	Northern Ireland Ireland Scotland Finland Other/ prefer not to say/ Unknown	1 (25%) - - - 4 (75%)	1 (<1%) 1 (<1%) - - 167 (99%)	5 (13%) 2 (5%) 1 (3%) 3 (8%) 27 (71%)

IV. DISCUSSION

This study provides an exemplar of the type of results that can be generated by analysing mental health chatbot data. Quantitative user event log analysis provides valuable insights into actual usage and allows better understanding of engagement and retention compared to conventional methods for app evaluation such as usability testing.

The chatbot was downloaded and used by a variety of individuals, across different ages, genders and countries. Only a small number of users (n=25) completed the WHO-5, and given the average score was indicative of poor mental wellbeing, it may be that only those with poor mental wellbeing were interested in completing this scale. Clustering revealed distinct characteristics of users within each of the groups. The largest cluster of chatbot users were individuals who downloaded ChatPal and only interacted with it for a short period of time. This is common with mental health apps and apps in general, as the majority of users do not progress to sustained long term use [21]. Lack of in person contact, workload/ time required, usability/ interface issues among many other contributing factors can result in high dropout rates in eHealth trials [22]. Overall usage was sporadic and user retention was poor, as none of the users progressed to long term use of the app. This is to be expected with 'in the wild' studies, as within other methodologies such as RCT participants will receive regular reminders or incentives to continue with the study. Strategies can be used to improve user retention such as increasing the amount of personalisation within the chatbot and prompting users to open the app, for example to try a new feature. A notification strategy that allows users to pick a day and time to receive prompts would be useful to encourage sustained use [23]. Personalised notifications have been shown to be effective in encouraging app use early on, while tailored insights encourage usage over time [24]. Alternatively, it may be the case that participants chose not to continue using the chatbot as they had a beneficial mental health outcome and got what they needed from the chatbot during their initial use.

As gender, age and country varied across clusters and given the majority of this information was not available, these characteristics were not useful for differentiating between clusters in this study. In the future and for other studies, knowing the demographics of those who did not proceed to sustained use of the chatbot would allow strategies to be developed to target these specific groups as a way of increasing engagement.

Given the end of a session was recorded infrequently, this indicates that people do not feel the need to say goodbye to a chatbot, where they would say goodbye to someone in a regular in person conversation. This also highlights that the end of a chatbot conversation is not the best time to present feedback to users. Thus, it may be best to ask for user feedback using questionnaires such as CUQ [19] randomly during a session. Alternatively, users could rate each individual conversation as a way of providing valuable feedback to app designers/ developers.

Creating a mental health chatbot is challenging. It is important to consider the ethical issues of human-computer conversations, such as the lack of true emotions and empathy and ensuring people do not over trust chatbot advice or recommendations. Another major challenge in the field is addressing the limitations of NLP. A recent review explored opinions of mental health chatbots and found that patients desired improvements to be made in NLU, in particular responding to unexpected input from users and providing high-quality variable responses.

A. Limitations

It was not possible to differentiate between 'in the wild' users and individuals who were involved in the project that downloaded and used the chatbot, which may have impacted the results. The majority of users did not continue using the chatbot long enough to provide their demographic information, which is another limitation of the study. In addition, there was no way to accurately measure when a session 'ended' as the vast majority of users did not formally 'exit' the app (select 'need to go' to the chatbot) which meant we could not study tenure. Finally, we were unable to calculate the total number of app users over the entire 8 month data collection period due to an error in labelling user IDs which was later corrected.

B. Future Work

Work is currently underway to develop new content and features in the chatbot and the next version of ChatPal will be multilingual. Future work will involve trialing the new version of ChatPal 'in the wild' and through inter-regional trials across rural areas in Europe. The trials will build on this initial study, measuring engagement, adoption and efficacy of the chatbot for supporting mental wellbeing in rural areas. Mixed method evaluation will be used to analyse user log data and other data collected through the chatbot. Outcome measures such as mental health scales will be recorded and can be correlated to app usage to assess effectiveness.

C. Conclusion

We recommend taking a responsible approach for mental health chatbot design which considers 1) what users want, 2) what mental health professionals will endorse and 3) what AI/ digital technologies can do well. As this minimum viable chatbot mainly included psychoeducational content and given that user retention was poor with this version of the chatbot, the next version of the chatbot will include more interactive features and dynamic content. This includes mood logging, mood visualisation and feedback, a gratitude diary, among other features based on user needs and what professionals will endorse. A key lesson learned from this study is the importance of adopting an iterative systems-centered design approach, where all stakeholders are involved and not just the end-user. This is crucial given that interventions and services can live in complex healthcare systems. In our experience, we found that it was important to meet only those user needs using technologies and features that are endorsed by healthcare professionals. This avoids creating technologies that users say they need but perhaps are met using suboptimal or 'potentially harmful' digital solutions that professionals would not endorse.

ACKNOWLEDGMENT

The authors would like to thank all the clients, participants, project members, supporters, and researchers at Ulster University, University of Eastern Finland, Norrbotten Association of Local Authorities, Region Norrbotten, Luleå University of Technology, NHS Western Isles, Action Mental Health and Munster Technological University for their involvement in this research.

REFERENCES

[1] S. Hoermann, K. L. McCabe, D. N. Milne, and R. A. Calvo, "Application of synchronous text-based dialogue systems in mental health interventions: Systematic review," Journal of Medical Internet Research, vol. 19, no. 8. Journal of Medical Internet Research, 01-Aug-2017.

[2] S. Provoost, H. M. Lau, J. Ruwaard, and H. Riper, "Embodied conversational agents in clinical psychology: A scoping review," Journal of Medical Internet Research, vol. 19, no. 5. Journal of Medical Internet Research, 01-May-2017.

[3] A. N. Vaidyam, H. Wisniewski, J. D. Halamka, M. S. Kashavan, and J. B. Torous, "Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape," Canadian Journal of Psychiatry, vol. 64, no. 7. SAGE Publications Inc., pp. 456–464, 01-Jul-2019.

[4] H. Gaffney, W. Mansell, and S. Tai, "Conversational Agents in the Treatment of Mental Health Problems: Mixed-Method Systematic Review," JMIR Ment. Heal., vol. 6, no. 10, p. e14166, 2019.

[5] K. H. Ly, A. M. Ly, and G. Andersson, "A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods," Internet Interv., vol. 10, pp. 39–46, Dec. 2017.

[6] R. A. Calvo, K. Dinakar, R. Picard, H. Christensen, and J. Torous, "Toward impactful collaborations on computing and mental health," J. Med. Internet Res., vol. 20, no. 2, Feb. 2018.

[7] T. Bird, W. Mansell, J. Wright, H. Gaffney, and S. Tai, "Manage Your Life Online: A Web-Based Randomized Controlled Trial Evaluating the Effectiveness of a Problem-Solving Intervention in a Student Sample," Behav. Cogn. Psychother., vol. 46, no. 5, pp. 570–582, Sep. 2018.

[8] S. Suganuma, D. Sakamoto, and H. Shimoyama, "An Embodied Conversational Agent for Unguided Internet-Based Cognitive Behavior Therapy in Preventative Mental Health: Feasibility and Acceptability Pilot Trial," JMIR Ment. Heal., vol. 5, no. 3, p. e10454, Jul. 2018.

[9] Y. T. Xiang et al., "Timely mental health care for the 2019 novel coronavirus outbreak is urgently needed," The Lancet Psychiatry, vol. 7, no. 3. Elsevier Ltd, pp. 228–229, 01-Mar-2020.

[10] A. S. Miner, L. Laranjo, and A. B. Kocaballi, "Chatbots in the fight against the COVID-19 pandemic," npj Digital Medicine, vol. 3, no. 1. Nature Research, pp. 1–4, 01-Dec-2020.

[11] A. A. Abd-alrazaq, M. Alajlani, A. A. Alalwan, B. M. Bewick, P. Gardner, and M. Househ, "An overview of the features of chatbots in mental health: A scoping review," Int. J. Med. Inform., vol. 132, no. June, p. 103978, 2019.

[12] M. E. P. Seligman, Flourish: A visionary new understanding of happiness and well-being. New York, NY: Simon and Schuster, 2011.

[13] S. J. Slavin, D. Schindler, J. T. Chibnall, G. Fendell, and M. Shoss, "PERMA: A model for institutional leadership and culture change," Acad. Med., vol. 87, no. 11, p. 1481, 2012.

[14] J. M. Norrish, P. Williams, M. O'Connor, and J. Robinson, "An applied framework for Positive Education," Int. J. Wellbeing, vol. 3, no. 2, pp. 147–161, 2013.

[15] S. Greer, D. Ramo, Y. J. Chang, M. Fu, J. Moskowitz, and J. Haritatos, "Use of the chatbot 'vivibot' to deliver positive psychology skills and promote well-being among young people after cancer treatment: Randomized controlled feasibility trial," JMIR mHealth uHealth, vol. 7, no. 10, pp. 1–13, 2019.

[16] C. Sweeney et al., "Can Chatbots Help Support a Person's Mental Health? Perceptions and Views from Mental Healthcare Professionals and Experts," ACM Trans. Comput. Healthc., vol. 2, no. 3, pp. 1–16, 2021.

[17] M. E. P. Seligman, T. A. Steen, N. Park, and C. Peterson, "Positive psychology progress: empirical validation of interventions.," Am. Psychol., vol. 60, no. 5, pp. 410–421, 2005.

[18] C. W. Topp, S. D. Østergaard, S. Søndergaard, and P. Bech, "The WHO-5 well-being index: A systematic review of the literature," Psychother. Psychosom., vol. 84, no. 3, pp. 167–176, Apr. 2015.

[19] S. Holmes, A. Moorhead, R. Bond, H. Zheng, V. Coates, and M. McTear, "Usability testing of a healthcare chatbot: Can we use conventional methods to assess conversational user interfaces?," in ECCE 2019 - Proceedings of the 31st European Conference on Cognitive Ergonomics: "'Design for Cognition,'" 2019, pp. 207–214.

[20] M. D. Mulvenna, R. R. Bond, A. Grigorash, S. O'Neill, and A. Ryan, "HILDA - A health interaction log data analysis workflow to aid understanding of usage patterns and behaviours," Proc. AISB Annu. Conv. 2018, 2018.

[21] A. Baumel, F. Muench, S. Edan, and J. M. Kane, "Objective User Engagement With Mental Health Apps: Systematic Search and Panel-Based Usage Analysis," J. Med. Internet Res., vol. 21, no. 9, p. e14567, 2019.

[22] G. Eysenbach, "The Law of Attrition," Journal of Medical Internet Research, vol. 7, no. 1. JMIR Publications Inc., p. e402, 31-Mar-2005.

[23] J. Melcher, E. Camacho, S. Lagan, and J. Torous, "College student engagement with mental health apps: analysis of barriers to sustained use," J. Am. Coll. Heal., vol. 0, no. 0, pp. 1–7, 2020.

[24] N. Bidargaddi, T. Pituch, H. Maaieh, C. Short, and V. Strecher, "Predicting which type of push notification content motivates users to engage in a self-monitoring app," Prev. Med. Reports, vol. 11, no. July, pp. 267–273, 2018.