
A Review on Acceptance of Conversational Agents in Health

Junhan Kim

University of Michigan
Ann Arbor, MI 48109, USA
ericjkh@umich.edu

Sun Young Park

University of Michigan
Ann Arbor, MI 48109, USA
sunypark@umich.edu

Lionel Robert

University of Michigan
Ann Arbor, MI 48109, USA
lprobert@umich.edu

Abstract

This paper presents a preliminary analysis of papers from the past review studies on conversational agents (CAs) in healthcare in terms of acceptance. We show a

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHI 2020 Extended Abstracts, April 24-30, Honolulu, HI, USA.

© 20w0 Copyright is held by the owner/author(s).

ACM ISBN 978-1-4503-6819-3/20/04.

DOI: <https://doi.org/10.1145/3334480>.

trend of user acceptance in CAs in healthcare by synthesizing the outcomes of different works on five main factors (i.e. satisfaction, ease of use, usefulness, the number of interactions, and the duration of interaction) of technological acceptance on CAs. We also describe directions for future research and design implications.

Author Keywords

Artificial intelligence; conversational agent; healthcare; virtual agent; wellbeing.

CSS Concepts

• **Human-centered computing~Human computer interaction (HCI)**

Introduction

Conversational agents (CAs, popularly known as chatbots) are now widely used across many fields [18,3,9]. Recent advances in artificial intelligence technology, widespread use of smartphones with high computing powers, and voice assistants such as Google Home or Alexa [1, 2] have allowed researchers to develop healthcare CAs coupled with machine learning capabilities in textual or auditory modalities as chatbots [18] or voice assistants [4]. Among those, the healthcare field in particular has adopted CAs as a patient care aid to support physical or mental health

[9] due to their convenience and patients' tendency to disclose personal information [22].

Due to the increased attention in CAs as a more accessible means of healthcare intervention, there have been works that provided an overview of the CAs supporting physical or mental illnesses [4,24]. However, there exists a gap in such review literature in that the past review papers have mostly concentrated on the design feature aspect of CAs in healthcare. We do not yet have a comprehensive review on the actual effects of CAs in health area. Therefore, it is important to analyze the studies according to actual outcomes of CAs in healthcare. We especially need to focus on acceptance (behavioral intention to use a system) to measure the influence CAs have on actual usage [11]. A recent integrated version of the technology

acceptance model by Wixom and Todd [28] states that satisfaction, ease of use, and usefulness are constructs that predict actual usage of a certain technology.

Therefore, we intend to provide the preliminary results of an ongoing review research trying to answer the following research questions: Do conversational agents work in healthcare in terms of acceptance? If so, what did the CAs in that literature try to accomplish? We answer these questions by conducting a review of the studies covered in review papers in this domain. We will present how much satisfaction, ease of use, and usefulness users feel in different works. Moreover, we will show how much users have actually encountered with the system. This work sets itself apart from other reviews in two ways. First, it tries to incorporate all CAs ranging from animated avatars to smartphone

CA Purpose	User acceptance							COUNT
	Satisfaction	Interaction	Duration	Enjoyment	Ease of Use	Usefulness	Intention to Use	
Data Collection	Rhee et al., 2014 [27] Harper et al., 2008 [11]	Kowatsch et al., 2017 ¹ [18]	Kowatsch et al., 2017 ¹ [18]	Kowatsch et al., 2017 ¹ [18] Kowatsch et al., 2017 ² [20]	Beveridge & Fox, 2006 [4] Kowatsch et al., 2017 ¹ [18]	Kowatsch et al., 2017 ² [20]	Kowatsch et al., 2017 ² [20]	10
Monitoring	Rhee et al., 2014 [27] Harper et al., 2008 [11] Kimani et al., 2016 [11]	Kowatsch et al., 2017 ¹ [18] Burton et al., 2016 [5]	Kowatsch et al., 2017 ¹ [18] Burton et al., 2016 [5]	Kowatsch et al., 2017 ¹ [18] Kowatsch et al., 2017 ² [20]	Kimani et al., 2016 [11] Kowatsch et al., 2017 ¹ [18]	Kowatsch et al., 2017 ² [20]	Kowatsch et al., 2017 ² [20]	13
Psychotherapy	Fitzpatrick et al., 2017 [12]	Fitzpatrick et al., 2017 [12], Ly et al., 2017 [21]						3
Decision Support		Burton et al., 2016 [5]	Burton et al., 2016 [5]		Beveridge & Fox, 2006 [4]			3
Education	Fitzpatrick et al., 2017 [12], Ireland et al., 2016 [16], Elmasri & Maeder, 2016 [10] Kimani et al., 2016 [11]	Fitzpatrick et al., 2017 [12] Ly et al., 2017 [21]	Crutzen et al., 2011 [7]		Crutzen et al., 2011 [7] Kimani et al., 2016 [11]	Crutzen et al., 2011 [7] Fernandez-Luque et al., 2018 [11]		11
Adherence support	Kimani et al., 2016 [11]	Hudlicka, 2013 [15]	Hudlicka, 2013 [15]		Kimani et al., 2016 [11]			4
Practice	Ireland et al., 2016 [16]	Hudlicka, 2013 [15]	Hudlicka, 2013 [15]	Cruz-Sandoval & Favela, 2017 [8]	Cruz-Sandoval & Favela, 2017 [8]			5
Clinical Interview	Elmasri & Maeder, 2016 [10]							1
Personal Assistance						Fernandez-Luque et al., 2018 [11]		1
COUNT	13	10	7	5	9	5	2	50

Table 1. This table shows the different literature with varying CA purpose and measured outcomes

chatbots. Second, this work includes not only random-controlled experiments, but also quasi-experiments in order to evaluate the effectiveness of CAs in different aspects and measures.

Method

We conducted a review of three systematic review papers [22, 25, 24]. These were searched from databases such as JMIR, AMIA, JMS, and ACM based on the queries, “conversational agents” “healthcare” and “review.” Among all the papers covered in these review works, we selected sixteen studies relevant to the topic of user acceptance on CAs. The relevance criterion was whether the researchers measured key predictors of acceptance according to the integrated model of acceptance: satisfaction, ease of use and usefulness. Moreover, the studies that measured the actual use of the system were also included.

As seen in Table 1, we classified the studies according to the CA’s purpose and the measurement of user acceptance. Criteria for the CA’s purpose was partially adopted from Laranjo et al. [22], but decision support and adherence support were added to accommodate other papers. We standardized different Likert scales of studies into a scale of 1 to 7 to reflect an adequate variability of the outcomes in Table 2. In Table 3, “Interaction” represents a number of times per week users interacted with the CA, and “Duration,” represents minutes of usage per usage.

Preliminary Findings

In this section, we summarize the preliminary findings from the review on satisfaction, ease of use, usefulness, the number of interaction and duration.

Satisfaction was measured in six studies. Two studies [16, 27] did not include a quantitative measure, but qualitatively measured the satisfaction of users from interviews. The six studies found that satisfaction towards the CA was high among participants, ranging from 5.95 to 6.17 out of 7.

Measures	Author, Year	Health Domain	CA Purpose	CA Type	Outcomes (1~7)
Satisfaction	Fitzpatrick et al., 2017	Depression	Psychotherapy, Education	Smartphone Chatbot	6.02
	Ireland et al., 2016	Language impairment	Education, Practice	Smartphone Chatbot	High (No scale)
	Rhee et al., 2014	Asthma	Data collection, Monitoring	Smartphone Chatbot	High (No scale)
	Harper et al., 2008	Diabetes	Data collection, Monitoring	Telephone	5.95
	Elmasri & Maeder, 2016	Addiction	Clinical Interview, Education	Smartphone Chatbot	6.17
	Kimani et al., 2016	Cardiovascular disorder	Education, Adherence support, Monitoring	Smartphone Chatbot	6.04
Ease of Use	Crutzen et al., 2011	Sexual health, Substance abuse	Education	2D Avatar	3.35
	Beveridge & Fox, 2006	Breast cancer	Data collection, Clinician Decision Support	Desktop CA	Moderate (No scale)
	Kimani et al., 2016	Cardiovascular disorder	Education, Adherence Support, Monitoring	Smartphone Chatbot	6.20
	Cruz-Sandoval & Favela, 2017	Dementia	Practice	Smartphone Chatbot + Robot	5.74
	Kowatsch et al., 2017	Obesity	Data collection, Monitoring	Smartphone Chatbot	6.70
	Cheng et al., 2018	Diabetes	Data collection, Monitoring	Voice Assistant	5.60
Usefulness	Crutzen et al., 2011	Sexual health, Substance abuse	Education	2D Avatar	3.95
	Kowatsch et al., 2017	Obesity	Data collection, Monitoring	Smartphone Chatbot	5.90
	Fernandez-Luque et al., 2018	Obesity	Education, Personal Assistance	Smartphone Chatbot	5.60

Table 2. Outcomes of factors contributing to acceptance

Ease of use ranged from 3.35 to 6.70 out of 7, with one study evaluating their CA as moderate. The exceptionally low score of the study by Crutzen et al. may stem from the discrepancy between free-typed questions and message database underlying the chatbot, thereby making the communication more difficult for the users.

The usefulness ranged from 3.95 to 5.90 out of 7. The relatively low score of the study by Crutzen et al. is possibly accounted for by the limited number of satisfactory answers from the chatbot, presenting itself as less useful for participants.

Measures	Author, Year	Health Domain	CA Purpose	CA Type	Outcomes (times/week)
Interaction	Fitzpatrick et al., 2017	Depression	Psychotherapy, Education	Smart-phone Chatbot	6.02
	Hudlicka, 2013	Mental health (mindfulness)	Education, Practice	2D Avatar	4.48
	Ly et al., 2017	Mental Well-being	Psychotherapy, Education	Smart-phone Chatbot	8.86
	Burton et al., 2016	Depression	Monitoring, Decision Support System	3D Animation	2.63
Measures	Authors, Year	Health Domain	CA Purpose	CA Type	Outcomes (min/use)
Duration	Hudlicka, 2013	Mental health (mindfulness)	Education, Practice	2D Avatar	19
	Crutzen et al., 2011	Sexual health, Substance abuse	Education	2D Avatar	3.95
	Burton et al., 2016	Depression	Monitoring, Decision Support System	3D Animation	12.76

Table 3. The number of interactions per week and the minutes of usage per interaction

Table 3 describes the comparison of actual usage in the CAs. Participants interacted with the CAs from 2.63 to 8.86 times. The variability in the number of interactions may be caused by a variety of factors such as directions given by the researchers or the differing types of CAs. We could also find a trend that mobility of the CAs through smartphone intervention and a psychotherapy-based intervention played a role in increasing the number of interactions between CAs and participants compared to immobile desktop-based CAs or CAs designed for practicing or monitoring.

Moreover, the duration of interactions per usage ranged from 3.95 minutes to 19 minutes. The CAs with longer duration contained a session with specific steps for the participants to follow, compared to the shorter study, which was more concentrated on simple question-answering procedures. The fact that the CA in the study by Crutzen et al. [7] had a shorter usage duration along with lower ease of use and usefulness score compared to others may suggest there is a possible correlation between these variables.

Implications for Design and Research

Overall, there seems to be a general trend towards satisfaction with CAs in healthcare. This can be seen from the rating score of satisfaction averaging about six out of seven. However, users seem to find CAs generally useful, but less than they are satisfied with them. It also has a wider variability than satisfaction level. This may be the result of differing natural language understanding capabilities of CAs. Users also find CAs in healthcare generally useful, but more so when they can be accessed through mobile and serve to handle obesity issues.

Through the analysis, we drew a few implications for future research and design of CAs. First, most of the CAs have been developed targeting the patient population, not the care providers. Benefits of CA for patient care is facilitating data collection by adding more types of patient-generated health data and monitoring and intervening patient's health-related behaviors. However, how these benefits are perceived by their clinicians and how they can be incorporated into the treatment process have not been examined yet. An integrated model of CAs mediating between patient and care provider will generate interesting discussions and provide a deeper insight into the behaviors and perceptions of users in the healthcare domain.

Second, most of the CAs have been rule-based, with a limited capacity to understand natural language. A future study that explores the capacity of chatbots with a large dialogue corpora dataset and how users interact with them will set the groundwork for forthcoming chatbots coupled with ML capabilities.

Finally, a deeper analysis of the user's perception towards CAs through qualitative methods is necessary, especially in terms of acceptance factors mentioned in this paper. It is important to understand how and why users find interactions with CAs satisfactory, easy to use and useful, which is a gap in current literature on this topic. Furthermore, a closer examination of each measurement will generate greater insight into users' perception. What are they exactly satisfied with? Is it the content? Is it the feature? What does it really mean to be satisfied? These are some of the questions that we would like to discuss in depth with other researchers at this workshop for CHI 2020. We hope to be able to share interesting ideas and insights at the workshop.

References

1. https://store.google.com/product/google_home
2. <https://www.amazon.com/all-new-amazon-echo-speaker-with-wifi-alexa-dark-charcoal/dp/B06XCM9LJ4>
3. Andreea I Niculescu, Kheng Hui Yeo, Luis F D'Haro, Seokhwan Kim, Ridong Jiang & Rafael E Banchs. 2014. Design and evaluation of a conversational agent for the touristic domain. *In Proceedings of APSIPA'14*: 1–10. doi:10.1109/APSIPA.2014.7041744
4. Beveridge, Martin, and John Fox. "Automatic generation of spoken dialogue from medical plans and ontologies." *Journal of biomedical informatics* 39, no. 5 (2006): 482-499.
5. Burton, Christopher, Aurora Szentagotai Tatar, Brian McKinstry, Colin Matheson, Silviu Matu, Ramona Moldovan, Michele Macnab et al. "Pilot randomised controlled trial of Help4Mood, an embodied virtual agent-based system to support treatment of depression." *Journal of telemedicine and telecare* 22, no. 6 (2016): 348-355.
6. Cheng, Amy, Vaishnavi Raghavaraju, Jayanth Kanugo, Yohanes P. Handrianto, and Yi Shang. "Development and evaluation of a healthy coping voice interface application using the Google home for elderly patients with type 2 diabetes." In *2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, pp. 1-5. IEEE, 2018.
7. Crutzen, Rik, Gjalt-Jorn Y. Peters, Sarah Dias Portugal, Erwin M. Fisser, and Jorne J. Grolleman. "An artificially intelligent chat agent that answers adolescents' questions related to sex, drugs, and alcohol: an exploratory study." *Journal of Adolescent Health* 48, no. 5 (2011): 514-519.
8. Cruz-Sandoval, Dagoberto, and Jesús Favela. "Semi-autonomous conversational robot to Deal

- with problematic behaviors from people with dementia." In *International conference on ubiquitous computing and ambient intelligence*, pp. 677-688. Springer, Cham, 2017.
9. Mauro Dragone, Thomas Holz, Brian R. Duffy, Gregory M.P. O'Hare. 2005. Social Situated Agents in Virtual, Real and Mixed Reality Environments. In *International Workshop on Intelligent Virtual Agents*, pp. 166-177. Springer, Berlin, Heidelberg, 2005.
 10. Elmasri, Danielle, and Anthony Maeder. "A conversational agent for an online mental health intervention." In *International Conference on Brain Informatics*, pp. 243-251. Springer, Cham, 2016.
 11. Fernandez-Luque, Luis, Abdelkader Lattab, Santiago Hors, and M. O. H. A. M. E. D. Ahmedna. "Implementation and feasibility study of a tailored health education bot in Telegram for mothers of children with obesity and overweight." In *Qatar Foundation Annual Research Conference Proceedings Volume 2018 Issue 2*, vol. 2018, no. 2, p. HBP506. Hamad bin Khalifa University Press (HBKU Press), 2018.
 12. Fitzpatrick, Kathleen Kara, Alison Darcy, and Molly Vierhile. "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, no. 2 (2017): e19.
 13. Harper, Roy, Peter Nicholl, Michael McTear, Jonathan Wallace, Lesley-Ann Black, and Patricia Kearney. "Automated phone capture of diabetes patients readings with consultant monitoring via the web." In *15th Annual IEEE International Conference and Workshop on the Engineering of Computer Based Systems (ecbs 2008)*, pp. 219-226. IEEE, 2008.
 14. Holden, Richard J., and Ben-Tzion Karsh. "The technology acceptance model: its past and its future in health care." *Journal of biomedical informatics* 43, no. 1 (2010): 159-172.
 15. Hudlicka, Eva. "Virtual training and coaching of health behavior: Example from mindfulness meditation training." *Patient education and counseling* 92, no. 2 (2013): 160-166.
 16. Ireland, David, Christina Atay, Jacki Liddle, Dana Bradford, Helen Lee, Olivia Rushin, Thomas Mullins et al. "Hello Harlie: enabling speech monitoring through chat-bot conversations." In *Digital Health Innovation for Consumers, Clinicians, Connectivity and Community-Selected Papers from the 24th Australian National Health Informatics Conference, HIC 2016, Melbourne, Australia, July 2016.*, vol. 227, pp. 55-60. IOS Press Ebooks, 2016.
 17. Kimani, Everlyne, Timothy Bickmore, Ha Trinh, Lazlo Ring, Michael K. Paasche-Orlow, and Jared W. Magnani. "A smartphone-based virtual agent for atrial fibrillation education and counseling." In *International conference on intelligent virtual agents*, pp. 120-127. Springer, Cham, 2016.
 18. S., Kopp, L., Gesellensetter, N.C., Krämer, & I., Wachsmuth. (2005). A Conversational Agent as Museum Guide – Design and Evaluation of a Real-World Application. In Gratch J. Panayiotopoulos T., Aylett R., Ballin D., Olivier P., Rist T. (eds) (Ed.), *Intelligent Virtual Agents*: Springer, Berlin, Heidelberg.
 19. Kowatsch, Tobias, Marcia Nißen, Chen-Hsuan Iris Shih, Dominik Rügger, Dirk Volland, Andreas Filler, Florian Künzler et al. "Text-based healthcare chatbots supporting patient and health professional teams: preliminary results of a randomized controlled trial on childhood obesity." (2017).
 20. Kowatsch, Tobias, Dirk Volland, Iris Shih, Dominik Rügger, Florian Künzler, Filipe Barata, Andreas Filler et al. "Design and evaluation of a mobile chat app for the open source behavioral health intervention platform MobileCoach."

In *International Conference on Design Science Research in Information System and Technology*, pp. 485-489. Springer, Cham, 2017.

21. Ly, Kien Hoa, Ann-Marie Ly, and Gerhard Andersson. "A fully automated conversational agent for promoting mental well-being: a pilot RCT using mixed methods." *Internet interventions* 10 (2017): 39-46.
22. Laranjo, Liliana, Adam G. Dunn, Huong Ly Tong, Ahmet Baki Kocaballi, Jessica Chen, Rabia Bashir, Didi Surian et al. "Conversational agents in healthcare: a systematic review." *Journal of the American Medical Informatics Association* 25, no. 9 (2018): 1248-1258.
23. 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.
24. Martínez-Miranda, Juan. "Embodied conversational agents for the detection and prevention of suicidal behaviour: current applications and open challenges." *Journal of medical systems* 41, no. 9 (2017): 135.
25. Pereira, Juanan, and Óscar Díaz. "Using health chatbots for behavior change: a mapping study." *Journal of Medical Systems* 43, no. 5 (2019): 135.
26. Philip, Pierre, Jean-Arthur Micoulaud-Franchi, Patricia Sagaspe, Etienne De Sevin, Jérôme Olive, Stéphanie Bioulac, and Alain Sauteraud. "Virtual human as a new diagnostic tool, a proof of concept study in the field of major depressive disorders." *Scientific reports* 7, no. 1 (2017): 1-7.
27. Rhee, Hyekyun, James Allen, Jennifer Mammen, and Mary Swift. "Mobile phone-based asthma self-management aid for adolescents (mASMAA): a feasibility study." *Patient preference and adherence* 8 (2014): 63.
28. Wixom, Barbara H., and Peter A. Todd. "A theoretical integration of user satisfaction and technology acceptance." *Information systems research* 16, no. 1 (2005): 85-102.