
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

In this paper, we present a preliminary analysis of past review studies on conversational agents (CAs) in health care in terms of user acceptance. We show a trend of

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>.

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

Author Keywords

Artificial intelligence; conversational agent; health care; 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 [3, 9, 18]. Recent advances in artificial intelligence technology, widespread use of smartphones with high computing powers, and voice assistants such as Google Home and Alexa [1, 2] have allowed researchers to develop health care CAs coupled with machine learning capabilities in textual or auditory modalities as chatbots [18] or voice assistants [4]. Among these, the health care field in particular has adopted CAs as a patient care aid to support physical or mental health [9] because of their convenience and patients' tendency to disclose personal information [22].

Because of the increased attention on CAs as a more accessible means of health care intervention, some researchers have provided an overview of CA support in physical and 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 features of CAs in health care. We do not yet have a comprehensive review of the actual effectiveness of CAs in health care. Therefore, it is important to analyze the studies according to actual outcomes of CAs in health care. 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

Therefore, we provide the preliminary results of an ongoing review project to answer the following research questions: Do conversational agents work in health care in terms of acceptance? If so, what did the CAs in that literature try to accomplish? We answered 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 incorporates all CAs ranging from animated avatars to smartphone chatbots. Second, this work includes not only random-controlled experiments,

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 [13]	Kowatsch et al., 2017 ² [19]	Kowatsch et al., 2017 ² [19]	Kowatsch et al., 2017 ² [19] Kowatsch et al., 2017 ² [20]	Beveridge & Fox, 2006 [4] Kowatsch et al., 2017 ² [19]	Kowatsch et al., 2017 ² [20]	Kowatsch et al., 2017 ² [20]	10
Monitoring	Rhee et al., 2014 [27] Harper et al., 2008 [13] Kimani et al., 2016 [17]	Kowatsch et al., 2017 ² [19] Burton et al., 2016 [5]	Kowatsch et al., 2017 ² [19] Burton et al., 2016 [5]	Kowatsch et al., 2017 ² [19] Kowatsch et al., 2017 ² [20]	Kimani et al., 2016 [17] Kowatsch et al., 2017 ² [19]	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 [17]	Fitzpatrick et al., 2017 [12] Ly et al., 2017 [21]	Crutzen et al., 2011 [7]		Crutzen et al., 2011 [7] Kimani et al., 2016 [17]	Crutzen et al., 2011 [7] Fernandez-Luque et al., 2018 [11]		11
Adherence support	Kimani et al., 2016 [17]	Hudlicka, 2013 [15]	Hudlicka, 2013 [15]		Kimani et al., 2016 [17]			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	51

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

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

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, 24, 25]. 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 16 studies relevant to the topic of user acceptance of CAs. The relevant 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 were 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 the number of times per week users interacted with the CA, and “Duration,” represents minutes of usage per interaction.

PRELIMINARY FINDINGS

In this section, we summarize the preliminary findings from the review on satisfaction, ease of use, usefulness, the number of interactions and the duration of interactions. 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 with 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 of conversational agents (CAs) on a scale of 1 (low) to 7 (high)

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

Measures	Author, Year	Health Domain	CA Purpose	CA Type	Outcomes (times/week)
Number of interactions	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 of interactions	Hudlicka, 2013	Mental health (mindfulness)	Education, Practice	2D avatar	19.00
	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 with conversation agent (CA)

Table 3 describes the comparison of actual usage of the CAs. Participants interacted with the CAs from 2.63 to 8.86 times. The variability in the number of interactions might 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, as compared to immobile desktop-based CAs or CAs designed for practicing or monitoring.

Moreover, the duration of usage per interaction ranged from 3.95 minutes to 19 minutes. The CAs with longer study 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 might suggest a possible correlation between these variables.

IMPLICATIONS FOR DESIGN AND RESEARCH

Overall, there seems to be a general trend toward satisfaction with CAs in health care. This can be seen from the rating score of satisfaction averaging about 6 out of 7. However, ratings on usefulness were lower. Usefulness ratings also had a wider variability than satisfaction levels. This might be the result of differing natural language understanding capabilities of CAs. Users also found CAs in health care to be more useful when they could be accessed through mobile devices and help them in handling 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 to assist the patient population, not the care providers. Benefits of CAs for patient care include (1) facilitating data collection by adding more types of patient-generated health data and (2) monitoring and intervening in patients' health-related behaviors. However, the ways that these benefits are perceived by clinicians and how they can be incorporated into the treatment process have not been examined. An integrated model of CAs mediating between patients and care providers could generate interesting discussions and provide deeper insight into the behaviors and perceptions of users in the health care 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 machine learning capabilities.

Finally, a deeper analysis of the user's perception of 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 and useful, which is a gap in current literature on this topic. Furthermore, a closer examination of each measurement would generate greater insight into users' perceptions. 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] Google Home website. 2020. Retrieved March 2, 2020, from https://store.google.com/product/google_home
- [2] Echo on Amazon website. 2020. Retrieved March 2, 2020, from <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, and 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] Martin Beveridge and John Fox. 2006. Automatic generation of spoken dialogue from medical plans and ontologies. *Journal of Biomedical Informatics* 39, 5 (2006), 482-499.
- [5] Christopher Burton, Aurora Szentagotai Tatar, Brian McKinstry, Colin Matheson, Silviu Matu, Ramona Moldovan, Michele Macnab et al. 2016. Pilot randomised controlled trial of Help4Mood, an embodied virtual agent-based system to support treatment of depression. *Journal of Telemedicine and Telecare* 22, 6 (2016), 348-355.
- [6] Amy Cheng, Vaishnavi Raghavaraju, Jayanth Kanugo, Yohanes P. Handrianto, and Yi Shang. 2018. 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] Rik Crutzen, Gjalt-Jorn Y. Peters, Sarah Dias Portugal, Erwin M. Fisser, and Jorne J. Grolleman.

2011. An artificially intelligent chat agent that answers adolescents' questions related to sex, drugs, and alcohol: an exploratory study. *Journal of Adolescent Health* 48, 5 (2011), 514-519.
- [8] Dagoberto Cruz-Sandoval and Jesús Favela. 2017. Semi-autonomous conversational robot to deal with problematic behaviors from people with dementia. In *International Conference on Ubiquitous Computing and Ambient Intelligence*. Springer, Cham, 677-688.
- [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*. Springer, Berlin, Heidelberg, 166-177.
- [10] Danielle Elmasri and Anthony Maeder. 2016. A conversational agent for an online mental health intervention. In *International Conference on Brain Informatics*. Springer, Cham, 243-251.
- [11] Luis Fernandez-Luque, Abdelkader Lattab, Santiago Hors, and M. O. H. A. M. E. D. Ahmedna. 2018. 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 2018*, 2, HBPD506. Hamad bin Khalifa University Press (HBKU Press).
- [12] Kathleen Kara Fitzpatrick, Alison Darcy, and Molly Vierhile. 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 (2017), e19.
- [13] Roy Harper, Peter Nicholl, Michael McTear, Jonathan Wallace, Lesley-Ann Black, and Patricia Kearney. 2008. 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)*, 219-226. IEEE.
- [14] Richard J. Holden and Ben-Tzion Karsh. 2010. The technology acceptance model: its past and its future in health care. *Journal of Biomedical Informatics* 43, 1 (2010), 159-172.
- [15] Eva Hudlicka. 2013. Virtual training and coaching of health behavior: example from mindfulness meditation training. *Patient Education and Counseling* 92, 2 (2013), 160-166.
- [16] David Ireland, Christina Atay, Jacki Liddle, Dana Bradford, Helen Lee, Olivia Rushin, Thomas Mullins et al. 2016. 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, 227, 55-60. IOS Press Ebooks.
- [17] Everlyne Kimani, Timothy Bickmore, Ha Trinh, Lazlo Ring, Michael K. Paasche-Orlow, and Jared W. Magnani. 2016. A smartphone-based virtual agent for atrial fibrillation education and counseling. In *International Conference on Intelligent Virtual Agents*, 120-127. Springer, Cham.
- [18] S. Kopp, L. Gesellensetter, N.C. Krämer, and I. Wachsmuth. 2005. A conversational agent as museum guide – design and evaluation of a real-world application. In *Intelligent Virtual Agents*, J. Gratch, T. Panayiotopoulos, R. Aylett, D. Ballin, P. Olivier, and T. Rist (eds.). Springer, Berlin, Heidelberg.
- [19] Tobias Kowatsch, Marcia Nißen, Chen-Hsuan Iris Shih, Dominik Rügger, Dirk Volland, Andreas Filler, Florian Künzler et al. 2017. Text-based healthcare chatbots supporting patient and health professional teams: preliminary results of a randomized controlled trial on childhood obesity. In *Persuasive*

Embodied Agents for Behavior Change (PEACH 2017) workshop, Stockholm, Sweden.

- [20] Tobias Kowatsch, Dirk Volland, Iris Shih, Dominik Rügger, Florian Künzler, Filipe Barata, Andreas Filler et al. 2017. 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 Systems and Technology*. Springer, Cham, pp. 485-489.
- [21] Kien Hoa Ly, Ann-Marie Ly, and Gerhard Andersson. 2017. A fully automated conversational agent for promoting mental well-being: a pilot RCT using mixed methods. *Internet Interventions* 10 (2017), 39-46.
- [22] Liliana Laranjo, Adam G. Dunn, Huong Ly Tong, Ahmet Baki Kocaballi, Jessica Chen, Rabia Bashir, Didi Surian et al. 2018. Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association* 25, 9 (2018), 1248-1258.
- [23] G. M. Lucas, J. Gratch, A. King, and L. P. Morency, 2014. It's only a computer: virtual humans increase willingness to disclose. *Computers in Human Behavior*, 37, 94-100.
- [24] Juan Martínez-Miranda. 2017. Embodied conversational agents for the detection and prevention of suicidal behaviour: current applications and open challenges. *Journal of Medical Systems* 41, 9 (2017), 135.
- [25] Juanan Pereira and Óscar Díaz. 2019. Using health chatbots for behavior change: a mapping study. *Journal of Medical Systems* 43, 5 (2019), 135.
- [26] Pierre Philip, Jean-Arthur Micoulaud-Franchi, Patricia Sagaspe, Etienne De Sevin, Jérôme Olive, Stéphanie Bioulac, and Alain Sauteraud. 2017. Virtual human as a new diagnostic tool, a proof of concept study in the field of major depressive disorders." *Scientific Reports* 7, 1 (2017), 1-7.
- [27] Hyekyun Rhee, James Allen, Jennifer Mammen, and Mary Swift. 2014. Mobile phone-based asthma self-management aid for adolescents (mASMAA): a feasibility study. *Patient Preference and Adherence* 8 (2014), 63.
- [28] Barbara H. Wixom and Peter A. Todd. 2005. A theoretical integration of user satisfaction and technology acceptance. *Information Systems Research* 16, 1 (2005), 85-102.