The Effects of Data Visualization on User Perceptions of a Health Chatbot

Ha-Kyung Kong[*]
Seattle University
Seattle, Washington

Ruijia Zhu[§]

Audrey Zhixin Lu[§]
[§]University of Washington
Seattle, Washington

Mingao He[§]

Weichao Yuwen[†]
University of Washington
Tacoma, Washington

Abstract

Chatbots are increasingly used in healthcare as an accessible and scalable means to communicate information. While visualization can provide summaries of health data, their impact on the comprehension, perceived usefulness, and trustworthiness of chatbots is yet unknown. We examined the effects of visualization (intervention group) of health data compared to text-only representation (control group) among 96 family caregivers in a randomized controlled trial. The results showed that visualizing the relationship between symptom levels (e.g. stress) and health solutions had a significant impact on the comprehension of the health trends and the perceived effectiveness of the solution. While there was an increased trust level in the visualization group, the difference was not statistically significant. The qualitative feedback aligned with the quantitative results. We discussed future directions in leveraging the synergistic effects of visualization and chatbots to increase health literacy and encourage technology adoption.

Index Terms: Human-centered computing—Visualization—Empirical studies in HCI—; Applied computing—Health Informatics—

1 Introduction

Chatbots are becoming increasingly common in healthcare domains [11], offering services such as symptom monitoring [20] and information collection [2]. While chatbots could provide an accessible and scalable way of communicating and evaluating health information, patients' acceptance of chatbots depends on their comprehensibility, perceived usefulness, and trustworthiness [3]. As many chatbots collect numeric data from the patients (e.g., weight, pain ratings), data visualizations offer a natural way to present and summarize the information. However, the impact of visualizations on the use and perception of chatbots has been under-explored. While studies have shown that visualizing data could increase viewers' comprehension of the information [5, 19], a recent study in healthcare has found a contradicting result where the comprehension rate was higher in the text-only condition than in the line graph condition [21]. Furthermore, no study has measured the effect of visualizations in health chatbots on users' trust in the technology to our knowledge. We bridge this gap through a randomized controlled trial (RCT) examining the effect of visualizations in health chatbots on comprehension, perceived usefulness, and trust.

We approached the problem by applying data visualization on a specific healthcare chatbot, COCO (Care for Caregivers Online) for family caregivers [9, 15]. COCO provides family caregivers – particularly those with limited resources and high stress and burnout – with on-demand caregiving support and interactive self- and family-management skill development. Although technology has the potential to provide cost-effective and accessible support for family caregivers, only half of the caregivers currently use technological tools for caregiving with the most common purpose of financial tracking, even though the number one need for caregivers is related to managing their own health [23]. To promote the acceptance of COCO, it is essential that COCO presents information in a way that is comprehensible, useful, and trustworthy. Thus, we examined whether presenting health information in a visualization format rather than a text format affects viewers' comprehension, perceived usefulness, and trust and acceptance of the information.

[*]e-mail: hkong@seattleu.edu
[†]e-mail: wyuwen@uw.edu

2 Related Work

Studies have shown that effective visualizations of personal health data allow users without any knowledge background to understand their overall conditions and lifestyle risk factors [14]. In contrast, a number of studies [4,6,21] have reported the negative effect of health data visualization on comprehension such as misinterpretations [6] and lower comprehension levels compared to visual analogies [21]. Given the mixed results on the impact of visualization on health data comprehension, Grossman et al. called out for future work to explore the effect of different visualizations on comprehension. Furthermore, no existing work has measured the impact of visualization on people's perception and anticipated use of a *health* chatbot to our knowledge. Hearst and Tory's study on the impact of visualization on general conversational interfaces found strong individual differences in people's preference for seeing visualizations in chatbots, with 41% of the participants preferring the text-only interface. While studies on health chatbots have included visualizations as a potential feature of the chatbot [8, 24], visualizations were not the focus of their studies. Our work fills this gap on the impact of visualization on comprehension and perception of a health chatbot.

Establishing trustworthiness is crucial in healthcare as patients' trust in healthcare providers directly affect their behavior to seek medical care and adhere to treatment regimens [25]. As public distrusts toward data collecting agencies is growing, [1], one of the biggest reasons patients are resistant to using a health chatbot is because they "do not trust enough in chatbots to not being afraid of regretting using a chatbot for medication in the future" [16]. Wang and Siau define trust in the domain of healthcare chatbot as "the willingness of users to provide confidential information, accept the recommendations, and follow the suggestions" [26]. Prior works showed that people's willingness to share electronic health records (EHR) is largely impacted by their attitudes toward EHR, the level of control they have over the data [10, 27], and direct benefits from sharing data such as "tailored and personalized data analysis, integrated view, feedback, and others" [27]. As visualizations are personalized feedback based on shared data, we hypothesized that the visual presentation would increase users' willingness to share data. We also used people's trust in health recommendations as the second measurement of trust in the chatbot according to Wang and Siau's definition of trust.

3 Methodology

We investigated the research questions through a non-blinded RCT where caregivers in the visualization group shared their perceptions based on information in a visual format, and the control group received the same information in a text format. Mix-methods evaluations with both quantitative ratings and qualitative open-ended questions were used in a survey format. In this section, we describe the participants, intervention materials, and study procedures.

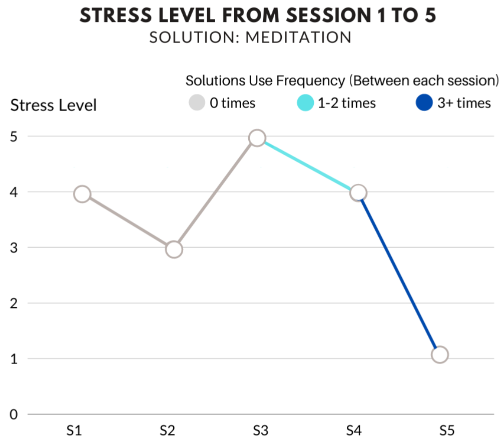


Figure 1: A visualization of stress level and solution use frequency used in the study. It shows a decreasing trend in the stress level starting from Session 3.

3.1 Participants

Participants were self-identified family caregivers in the United States. The inclusion criteria for participants were a) over the age of 18, b) currently taking care of a person with a chronic health condition, and c) can read and write in English. Caregivers were recruited through specialty care clinics (e.g., autism centers), social media and caregiving websites (e.g., caregiver support groups, Family Caregiver Alliance website), snowballing, and community email lists. A total of 96 family caregivers participated in the study.

3.2 Intervention materials

We created a data set in two different formats: text description (control) and visualization (intervention) that conveyed the same information. We used fabricated information about a caregiver's stress levels over five chat therapy sessions and how many times the caregiver had carried out the recommended health solution after each session. The text description was "Scenario: COCO worked with you through 5 therapy sessions over three weeks to decrease your stress level, and you chose a health solution (meditation) to practice regularly. Between Sessions 3, 4, and 5, you tried meditating a few times. COCO shared a summary of your reported stress levels from Session 1 to Session 5, which were: 4, 3, 5, 4, and 1."

The visual representation was designed to be parallel in terms of the information conveyed, as shown in Figure 1. The visualization was iteratively developed based on a series of pilot tests. Major changes made in the process include mapping solution usage frequency on the line color instead of the circle color, only showing one solution per visualization, simplifying the labels, and specifying the solution (e.g., meditation) in the heading. The trends conveyed in the two presentation formats (text and visual) were the same; the stress level started high in Session 1, reached the peak level in Session 3, and decreased afterward when the user started using the solution. The representations only differed on the breakdown of solution use frequency (i.e., 0 times, 1-2 times, 3+ times) in the visual representation. The solution frequency information was excluded from the text description as it required participants to keep track of the solution frequency in addition to the solution starting point, which could create an information overload.

3.3 Study Procedure

This RCT was part of a larger study approved by the University's Internal Review Board. All caregivers provided electronic consent for participation and then interacted with COCO daily through text-based dialog about their health using their mobile phones for two weeks. At the end of the two weeks, participants were invited to participate in the current study. We used a computer program to generate simple randomization group allocations prior to the study. Participants were randomly assigned to a condition based on the pre-assigned allocation and sent an electronic link via email with the corresponding study survey. The survey took approximately 10-15 minutes to complete. Participants were provided with a $10 e-gift card upon completion of the survey. Both groups were asked the same set of questions to assess: 1) the comprehension of the provided information, 2) the perceived usefulness of the information presented as part of COCO, 3) the trustworthiness of the information presented, and 4) the users' acceptance of COCO.

Comprehension. Two multiple-choice questions were about interpreting the information, such as detecting peaks, evaluating trends, and identifying correlations (e.g., "During which session did you report the highest stress level?"). Two additional questions assessed comprehension: "I believe the solution was effective in improving my health" and "I think it will be easy to evaluate the effectiveness of the solutions suggested by COCO with the information" with Likert scale ratings (1 = strongly disagree, 5 = strongly agree).

Usefulness. Three Likert scale questions measured the perceived usefulness of the provided information in three aspects: evaluating, tracking, and communicating information (e.g., "The information will be helpful for me to review my progress"). We asked follow-up open-ended questions to elicit an explanation of the ratings.

Trust and acceptance. Three sets of questions measured the perceived trustworthiness in three aspects: confidence in the information provided, trust in the health recommendations, and willingness to share health data with the system with the 5-point Likert scale. Similar follow-up questions were asked to elicit an explanation of the ratings. Acceptance was measured with two questions ("I would use COCO" and "I would recommend COCO to others").

Qualitative Feedback. In addition to the open-ended questions, another question was asked in the end: "What could we do to make these graphs more useful?" All qualitative data were analyzed by one researcher using deductive content analysis [7], coding into the three pre-determined themes: comprehension, usefulness, and trust. In each theme, data were coded into benefits, challenges, and suggestions. A separate researcher reviewed the coding, and discrepancies were discussed until 100% consensus was reached.

4 Results

A total of 96 family caregivers (52 in the visualization group and 44 in the control group) aged from 26 to 81 years participated in the study (mean age 40 years). For the people they cared for, common health conditions include autism, asthma, cancer, diabetes, and memory loss. The majority of participants were female (93%), working full-time (71%), and had an annual household income of $80,000 or more (54%). The majority were white (70%) and 27% were Asian.

4.1 Quantitative Results

RQ1. The visual representation significantly improves information comprehension. A summary of the ratings is presented in the figures in the Appendix. We first analyzed the comprehension levels between the two conditions. The answers to two comprehension questions were converted to binary data (correct or incorrect) and combined to represent the number of questions answered correctly per condition. The result using a Chi-squared test showed a significant difference in comprehension: the participants in the visualization condition were much more likely to answer questions correctly (χ^2 = 19.346, df = 2, $p < 0.001$). More specifically, 84% of the participants who referred to a visualization answered both questions correctly compared to 40% in the control group. Next, the independent samples t-tests showed significant differences between the perceived effectiveness of the solution (t(93)=-3.448, p=.001) where the solution was rated as more effective in the visualization

	Benefits	Challenges	Suggestions
Compre-hension	**Control Group:** "It gave me concrete numbers to show what I rated my stress level and I could track over time to see if what I was doing is working." **Visualization Group:** "I'm a visual person and this representation better helps me understand the connection between the solution and its potential effect on reported stress levels."	**Control Group:** "The change in stress levels was only reported in a single week, rather than after a few additional sessions after the intervention had been deployed." **Visualization Group:** "I just realized I was reading the graph incorrectly, the stress level went up! This is not a good graphic. Try another one."	**Control Group:** "A graph, summary of some kind other than text after the fact would be useful (like fitbit does with sleep levels). Something I can visualize would help me." **Visualization Group:** "Instead of color may be the frequency of health solution can be shown using thickness of the lines 0 - dotted, 1-2 - think, 3+ thick. Its hard to recall the color and frequency mapping while reading the plot."
Useful-ness	**Control Group:** "To be able to track progress and name feelings over time would be very helpful, especially when feeling overwhelmed or stressed. You can look back and understand better why you felt how you did. This inspires giving oneself more patience and grace." **Visualization Group:** "Rather than feelings and emotions, it is based on data (that is based on feelings and emotions)."	**Control Group:** "It would be helpful, but this survey is hard to understand." **Visualization Group:** "I think it can show progress but I when overwhelmed with other things it might not make as much sense or take some time to understand."	**Control Group:** "Also, putting the date with the session number […] could help people see if stress was particularly bad on a specific date and they know what happened during that date, like they lost their job, then it could help bring important context to the graph that could help make it more meaningful." **Visualization Group:** "The graph is helpful to track feedings and the self-help practice but life is dynamic so it doesn't reflect whether meditation was the cause of stress reduction or other factors in life."
Trust	**Control Group:** "Not all recommendations are perfect for my situation, but I trust based on my reduction in stress from a 5 to a 1 that Coco will give me recommendations that can be beneficial." **Visualization Group:** "As long as Coco makes recommendations that are informed by data like these graphs, I would trust them." "Research based recommendations. It is based on research, so I trust science."	**Control Group:** "As presented Coco is giving me confusing and unclear information, which makes me distrust its ability to help me in a meaningful way." **Visualization Group:** "Each situation is unique and we are self-reporting through an app so there is a lot of room for error and/or glossing over the true answers to each prompt."	**Control Group:** "I don't know anything about what it's basing its recommendations on." "If recommendations are accompanied by supporting articles/documents I would trust them even more." **Visualization Group:** "I don't know the people behind Coco nor do I understand their motivations. Knowing these may help me gain more trust, although it already seems quite beneficent."

Table 1: A summary of qualitative results illustrated through quotes on the benefits, challenges, and suggestions on comprehension, usefulness, and trust.

condition (mean=4.29, s.d.=0.94) compared to the control condition (mean=3.57, s.d.=1.1). This difference could have resulted from the visual representation of the decrease in stress between Sessions 3 and 5, as this trend is conveyed more clearly in the visualization compared to the textual list of numbers. Accordingly, results showed that the participants found it significantly easier to rate the effectiveness of a solution with a visualization (mean=4.0, s.d.=0.97) than with a text description (mean=3.43, s.d.=1.26) (t(79.79)=-2.435, p=.017). Overall, we found a significant difference between the conditions for every question on comprehension.

RQ2. The visual representation is perceived as more useful for reviewing progress. Next, we examined the effect of visualization on the perceived usefulness of the system and the information provided. While people perceived COCO to be more useful for reviewing progress in the visualization condition (mean=4.24, s.d.=1.06) compared to the control condition (mean=3.80, s.d.=1.23), the difference was not statistically significant (t(93)=-1.864, p=.065). The ratings for COCO's potential to help track symptoms were similar across conditions (t(93)=-0.159, p=.874). However, people found the visual representation to be significantly more useful (mean=4.18, s.d.=1.07) as something that COCO could show compared to the textual representation (mean=3.70, s.d.=1.02) (t(93)=-2.184, p=.031).

RQ3. The representation format does not affect users' trust or acceptance of the chatbot. We found no difference between the conditions about people's willingness to share information (t(94)=-0.774, p=.44). The overall rating for the data sharing questions was much lower than other ratings. People rated their trust in the health recommendation higher in the visualization condition (mean=3.96, s.d.=1.04) compared to in the control condition (mean=3.64, s.d.=1.10), but the difference was not statistically significant (t(93)=-1.476, p=.143). On the other hand, participants were significantly more confident in understanding the information when the information was presented in a visualization format (mean=4.27, s.d.=1.15) rather than in a text-only format (mean=3.45, s.d.=1.30) (t(93)=-3.307, p=.001). Lastly, we did not find a significant difference in people's acceptance of COCO based on the representation formats (t(94)=-1.527, p=.130).

4.2 Qualitative Results

Qualitative findings elaborated on the quantitative results. Representative quotes from each domain are shown in Table 1. Participants from both groups found that the information helped them understand their stress levels. While many in the control group suggested visualizing the data, participants in the visualization group commented that the visualization facilitated the comprehension of the relationship and trend. There were many recommendations for improving the visualization, including clarifying and adding more legends, color usage, and adding accompanying text explanations. Both groups shared that the data is useful to track their health, while one participant in the control group said the difficulty in understanding made it not as helpful. Participants suggested making the visualizations more comprehensive and dynamic (e.g., visualizing additional stress-related factors) rather than simplifying relationships. A related suggestion was to expand the data types and metrics, such as linking to data from wearable devices. The results regarding trust shed light on factors that contribute to the perceived trustworthiness of the program, including how data were used, the content and trend shown in the visualization, and how health recommendations were made based on resources beyond the personal data (e.g., whether they were evidence-based recommendations), and the creator of the program (e.g., this is created by scientists at the university).

5 Discussion

The results indicate that visualizations are effective for increasing viewers' comprehension of information and more useful to include in a health chatbot than text-only presentation. However, it did not have a significant influence on the perceived trustworthiness of COCO. We discuss these results in more detail and present three areas for future work.

5.1 The Role of Visualization in Health Chatbots

While prior works have yielded mixed results on the effects of visualization on health data comprehension, our findings showed that visual presentation could increase caregivers' comprehension of the information. A feature to collect and display information is not sufficient to be useful in a health chatbot. Although participants in both groups agreed that COCO had the potential to track their symptoms, visualization was perceived to be more useful for reviewing progress. By facilitating users in *evaluating* information, visualization can support the chatbot's decision process to make health recommendations more transparent. For example, a visualization with an increasing trend in stress may illustrate the ineffectiveness of the current solution. With this knowledge, users may better understand why the chatbot is suggesting them to try a different solution. Nadarzynski et al.'s work on the acceptability of AI-led chatbot services in healthcare revealed a general lack of familiarity and understanding of health chatbots amongst participants [17]. Moreover, they were also hesitant about adopting chatbots as they were unable to understand the technological complexity. In light of their finding, our results on comprehension and usefulness demonstrate a promising role of visualization in health chatbots.

Despite the positive influence on comprehension and the perceived usefulness, the visualization by itself was not sufficient to increase trust. The ratings for the data sharing questions were much lower compared to other ratings, indicating people's resistance to sharing personal data with health technology where trust has not been established. The qualitative results showed that participants considered multiple factors when evaluating the system's trustworthiness. Two factors were data-based – the inclusion of data and the visible trends shown – which could be enhanced by the design of data visualization. Other factors are independent of data, such as the creator's credentials and external supporting sources. This behavior of seeking evidence beyond the data at hand is reassuring as prior works have raised concerns regarding unwavering trust in data and data visualization, which could lead to blind acceptance of biased information [12, 13]. Health chatbot creators should encourage users' requests for the aforementioned factors through design and incorporate multiple ways to proactively build trust. To increase trust through data-based factors, two participants mentioned expanding from self-report data to integrating data such as heart rate from wearable technology.

5.2 Future Work

In our study, we have explored how visualization can improve a chatbot. Interestingly, the results revealed that the benefit could be reciprocated as the chatbot environment has the potential to improve the understanding and usefulness of visualization. In this section, we discuss three directions for future work that leverage the synergistic effect of visualization and chatbot to scaffold personalized feedback and increase health literacy.

Chatbot as a Guide. In our study, no introductory guidance was provided to explain the visual mapping or how to interpret the visual patterns. In other words, the participants had to rely on the visualization itself to learn how to read it. While most participants found the visualization easy to understand, some participants noted that the labels or the axes were confusing. This challenge can be overcome with an introductory guide to the visualization in a dialog, such as explaining the basic visual mappings such as "each circle represents a therapy session" or how to interpret a visual trend. Alternatively, the chatbot could use a staged transition where the visualization is built up one component at a time. For example, only the axes are presented at first with a brief explanation, then the circles appear, then the lines, etc. giving the viewers more time to learn new information. In the future, we plan to explore ways that a chatbot could provide an interactive introduction to visualizations and test their effects on users' graph literacy.

Chatbot as a Narrator. A health chatbot provides an ideal environment to combine data, texts, and graphs to tell a cohesive story. Thus one direction of research is to explore various text-based methods for integrating contextual information into visualizations. For example, the description "The biggest drop in stress was shown after Session 4" could be embedded directly on the visualization, shown in tooltips when an element is pressed on or conveyed through the chatbot's message. A chatbot also enables the exploration of data and patterns through conversations in addition to interactive visualizations. Future work could explore how users engage in conversations to seek more context, which can be especially useful in a mobile environment [18], and how visual cues can be used to answer questions. For example, when a user presses on a session number on the chart, the chatbot could provide relevant contextual information from the conversation they had during that session, such as whether the user had mentioned a stressful event at work.

Chatbot for Personalizing Visualizations. Next, the chatbot environment allows for personalized visualizations through long-term interactions with users and the ease of communicating personal preferences in a dialog. We can start the visualization with basic information for first-time users and add more comprehensive and personalized information over time. The visualization can also be dynamically adapted to match user preferences through dialog. For example, the user could follow a strict, linear narrative early on and later engage in an open exploration of the content [22], such as the inclusion of external data (e.g., heart rate or sleep quality data from wearable devices) and removal of visual elements (e.g., legend) to create a visual estate for more contextual illustrations. Communicating comprehensive information enhances trust, which in turn encourages the use of health chatbots.

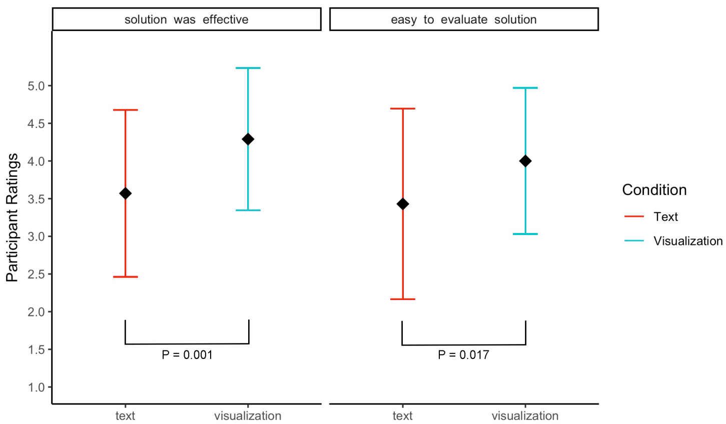
6 Limitations

Although we tried to recruit an inclusive sample through broad Internet recruitment and low-resource communities (e.g., families served by an autism center funded primarily through Medicare and Medicaid), the majority of the participants were from the communities near a large public university, where the population has relatively high income levels and employment rates. The participants' graph literacy levels were relatively high, but we did not collect educational level data to examine whether they were correlated. These limitations compromise the external validity of the findings. Future work with a more diverse population is required to determine the generalizability of our results. Furthermore, our study involved a single visualization type and the results might not be generalized to other visualization types of differing complexities. While we positioned the visualization as a part of a dialog system, the information was presented through a survey, not a chatbot environment. We intend to conduct a follow-up study with the discussed future directions delivered through a chatbot.

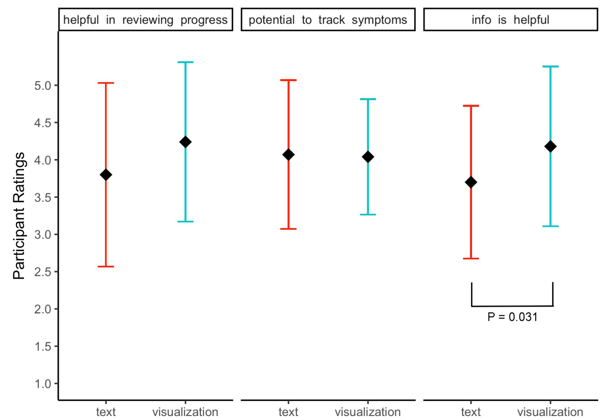
7 Conclusion

In this RCT, we tested the role of visualization on COCO, a health chatbot for family caregivers, to examine its effect on factors (comprehension, perceived usefulness, and trust) that could promote the use of the technology. The results showed that visual presentation could improve comprehension compared to text descriptions. The presentation format had a significant impact on the perceived effectiveness of the solution and participants strongly preferred the visual representation as information that COCO could show. While there was not a significant increase in trust or acceptance levels in the visualization group, the promising trends warrant additional examination. When evaluating the trustworthiness of COCO through visualization, caregivers considered various factors beyond their own health data such as research evidence and the credentials of the creators. The findings highlighted people's desire for more contextual data and guidance on interpreting the visualization. Based on these findings, we discussed ways to explore the role of chatbot to guide, narrate, and scaffold through personalized visualizations.

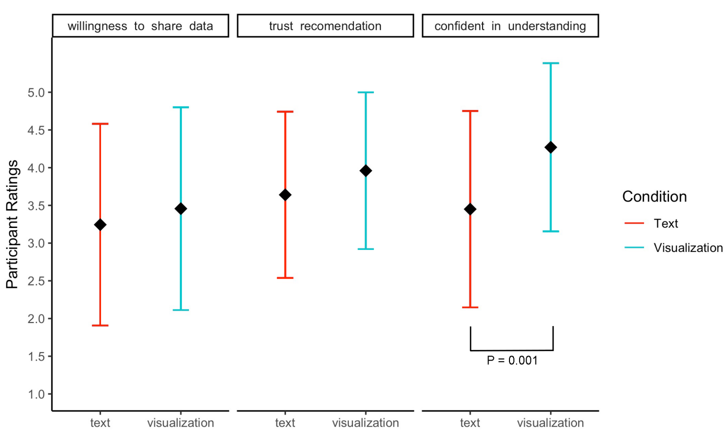
A Appendix



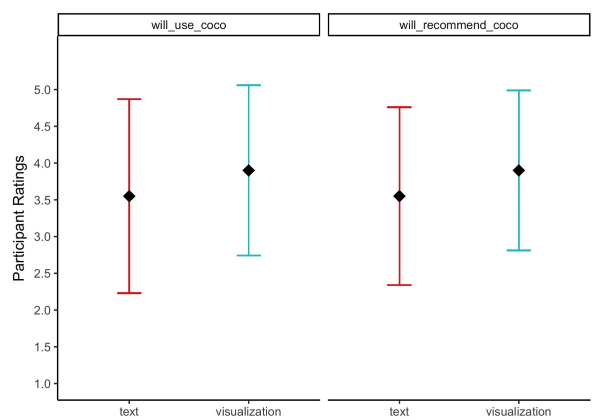
(a) Comprehension



(b) Usefulness



(c) Trust



(d) Acceptance

Figure 2: Four figures showing the user ratings for comprehension, usefulness, trust, and acceptance with mean values shown as black dots and the standard deviations shown as whiskers above and below the dots. Visualization condition had significantly higher ratings for all aspects of comprehension and one aspect of usefulness. Increasing trends are observed in the visualization condition for all variables in trust and acceptance, but only the confidence in understanding information was significantly improved.

REFERENCES

[1] B. Auxier, L. Rainie, M. Anderson, A. Perrin, M. Kumar, and E. Turner. Americans and privacy: Concerned, confused and feeling lack of control over their personal information. *Pew Research Center*.

[2] K. Denecke, S. Lutz Hochreutener, A. Pöpel, and R. May. Talking to ana: A mobile self-anamnesis application with conversational user interface. In *Proceedings of the 2018 international conference on digital health*, pp. 85–89, 2018.

[3] D. Dhagarra, M. Goswami, and G. Kumar. Impact of trust and privacy concerns on technology acceptance in healthcare: an indian perspective. *International journal of medical informatics*, 141:104164, 2020.

[4] L. V. Grossman, S. K. Feiner, E. G. Mitchell, and R. M. M. Creber. Leveraging patient-reported outcomes using data visualization. *Applied clinical informatics*, 9(03):565–575, 2018.

[5] T. D. Hendrix, J. H. Cross, S. Maghsoodloo, and M. L. McKinney. Do visualizations improve program comprehensibility? experiments with control structure diagrams for java. In *Proceedings of the thirty-first SIGCSE technical symposium on Computer science education*, pp. 382–386, 2000.

[6] Z. Hildon, D. Allwood, and N. Black. Making data more meaningful: patients' views of the format and content of quality indicators comparing health care providers. *Patient education and counseling*, 88(2):298–304, 2012.

[7] H.-F. Hsieh and S. E. Shannon. Three approaches to qualitative content analysis. *Qualitative health research*, 15(9):1277–1288, 2005.

[8] T.-H. Hwang, J. Lee, S.-M. Hyun, and K. Lee. Implementation of interactive healthcare advisor model using chatbot and visualization. In *2020 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 452–455. IEEE, 2020.

[9] W. R. Kearns, N. Kaura, M. Divina, C. Vo, D. Si, T. Ward, and W. Yuwen. A wizard-of-oz interface and persona-based methodology for collecting health counseling dialog. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–9, 2020.

[10] K. K. Kim, P. Sankar, M. D. Wilson, and S. C. Haynes. Factors affecting willingness to share electronic health data among california consumers. *BMC medical ethics*, 18(1):1–10, 2017.

[11] A. B. Kocaballi, J. C. Quiroz, L. Laranjo, D. Rezazadegan, R. Kocielnik, L. Clark, Q. V. Liao, S. Y. Park, R. J. Moore, and A. Miner. Conversational agents for health and wellbeing. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–8, 2020.

[12] H.-K. Kong, Z. Liu, and K. Karahalios. Frames and slants in titles of visualizations on controversial topics. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2018.

[13] H.-K. Kong, Z. Liu, and K. Karahalios. Trust and recall of information across varying degrees of title-visualization misalignment. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2019.

[14] A. Ledesma, H. Nieminen, P. Valve, M. Ermes, H. Jimison, and M. Pavel. The shape of health: A comparison of five alternative ways of visualizing personal health and wellbeing. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 7638–7641, 2015. doi: 10.1109/EMBC.2015.7320161

[15] Y. Liu, L. Wang, W. R. Kearns, L. Wagner, J. Raiti, Y. Wang, and W. Yuwen. Integrating a voice user interface into a virtual therapy platform. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–6, 2021.

[16] L. Müller, J. Mattke, C. Maier, and T. Weitzel. Conversational agents in healthcare: Using qca to explain patients' resistance to chatbots for medication. In A. Følstad, T. Araujo, S. Papadopoulos, E. L.-C. Law, O.-C. Granmo, E. Luger, and P. B. Brandtzaeg, eds., *Chatbot Research and Design*, pp. 3–18. Springer International Publishing, Cham, 2020.

[17] T. Nadarzynski, O. Miles, A. Cowie, and D. Ridge. Acceptability of artificial intelligence (ai)-led chatbot services in healthcare: A mixed-methods study. *Digital health*, 5:2055207619871808, 2019.

[18] S. H. Oh, S. J. Hur, and S.-H. Kim. Natural language processing model for data visualization interaction in chatbot environment. *KIPS Transactions on Computer and Communication Systems*, 9(11):281–290, 2020.

[19] J. Peltonen, K. Belorustceva, and T. Ruotsalo. Topic-relevance map: Visualization for improving search result comprehension. In *Proceedings of the 22nd international conference on intelligent user interfaces*, pp. 611–622, 2017.

[20] A. Piau, R. Crissey, D. Brechemier, L. Balardy, and F. Nourhashemi. A smartphone chatbot application to optimize monitoring of older patients with cancer. *International journal of medical informatics*, 128:18–23, 2019.

[21] M. Reading Turchioe, L. V. Grossman, A. C. Myers, D. Baik, P. Goyal, and R. M. Masterson Creber. Visual analogies, not graphs, increase patients' comprehension of changes in their health status. *Journal of the American Medical Informatics Association*, 27(5):677–689, 01 2020. doi: 10.1093/jamia/ocz217

[22] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16:1139–1148, 2010.

[23] The National Alliance for Caregiving. Caregiving in the U.S. 2020, 2020.

[24] P. Thongyoo, P. Anantapanya, P. Jamsri, and S. Chotipant. A personalized food recommendation chatbot system for diabetes patients. In *International Conference on Cooperative Design, Visualization and Engineering*, pp. 19–28. Springer, 2020.

[25] F. Trachtenberg, E. Dugan, and M. A. Hall. How patients' trust relates to their involvement in medical care. *Journal of Family Practice*, 54(4):344–354, 2005.

[26] W. Wang and K. Siau. Trust in health chatbots. 12 2018.

[27] A. Z. Woldaregay, A. Henriksen, D. Issom, G. Pfuhl, K. Sato, A. Richard, C. Lovis, E. Årsand, J. Rochat, and G. Hartvigsen. User expectations and willingness to share self-collected health data. *Studies in health technology and informatics*, 270:894–898, 2020.