Closing the Loop in Computer Agent/Patient Communication

While patient-centered communication supports patient self-care, providers rarely have enough time to consistently use this communication. Technology has potential to support patient-centered communication by elaborating provider communication, but frequently older adults prefer direct communication with providers. Conversational agents (CAs) may support provider communication with older adults by emulating best practices from face-to-face communication. We investigated older adults' response to a prototype CA communication system that served as a virtual provider and presented medication instructions using teachback, a recommended best practice that involves asking patients questions to 'close the communication loop'. Older adults were told how to take medications by a CA who used (or did not use) teachback, and then were interviewed. Older adults were open to interacting with the CA and thought it would help support self-care. They thought the CA was a more effective teacher when using teachback and that this interactive strategy helped them remember the instructions.

Introduction

Patient-centered communication helps patients understand and use concepts needed for self-care (Wagner et al., 2001). Unfortunately, providers have little time to consistently use this communication. Technology has the potential to support patient-centered communication by elaborating and reinforcing provider communication. However, older adults, especially those with less health literacy, tend to prefer face-to-face communication with providers (Medlock et al., 2015). Conversational agents (CAs) may support provider communication with older adults by emulating best practices from face-to-face communication. There is a growing literature exploring whether older adults are satisfied with and use CAs as support for self-care (Bickmore et al., 2010; Hanke et al., 2015; Strassman & Kramer, 2017), but this work has not focused on how older adults respond to specific CA-based interactive teaching strategies.

We investigated CA-based instruction about medication, focusing on teachback, a recommended best practice for educating patients (Paasche-Orlow et al., 2006). Teachback involves asking questions to ensure patients understand presented information. It allows patients to state in their own words what they need to know and do, therefore allowing providers to assess if patients understood their instructions (e.g., how to use an insulin injection). By 'closing the communication loop', teachback not only supports learning, but promotes patient/provider agreement about treatment goals and other aspects of health care (Schillinger et al., 2003). This patient/provider shared understanding has been found to predict patient satisfaction, knowledge, behavior (e.g., adherence), and outcomes (Street et al., 2009).

Teachback can improve patient learning by correcting misunderstanding, strengthen later retrieval of the information (the 'testing effect', Karpicke & Blunt, 2011), and perhaps by boosting motivation. However, while teachback improves patient memory (e.g., White et al., 2013), it is not routinely used because of limited provider time and training (AHRQ, 2019). CAs have potential to deliver teachback as needed. We investigated older adults' response to a prototype CA communication system that served as a virtual provider and used teachback. Responses were measured by open-ended and scaled questions, including the Agent Persona Inventory (API), which measures perceived informational usefulness and how relational the CA is (Ryu & Baylor, 2005).

We predicted that: 1) Older adults would be open to interacting with the CA; 2) CA-based teachback would increase the perceived usefulness of the CA as a virtual provider compared to a condition without teachback. More specifically, we predicted the CA would be perceived as more informationally useful (effective teacher). However, it would not be perceived as more affectively interactive (relational) when using teachback because the same CA was used in both conditions; 3) A secondary prediction was that teachback would improve memory for the instructions.

Method

Participants. 36 older adults participated, 18 in the teachback (TB) and 18 in a non-teachback (NTB) condition. All participants were community-dwelling older adults with no physical, cognitive or visual/auditory impairments that could limit the interaction with the CA. The groups did not differ in average age (TB = 73 years; NTB= 72 years), education (TB & NTB = 16.2 years), health literacy (5-point scale; Chew, Bradley & Boyko, 2004) (TB = 3.6; NTB = 3.7), or self-reported overall health condition (7-point scale; 1-very unhealthy/7-very healthy; TB = 5.3; NTB=5.6; $p>.10$ for all comparisons).

Procedure and Materials. A CA (developed using the Virtual Human Toolkit; https://vhtoolkit.ict.usc.edu/; Hartholt et al., 2013) told older adults how to take two medications for Type II diabetes. A female CA was used (Figure 1) because we earlier found that older and younger adults tend to prefer receiving medication instructions from female human-looking CAs rather than male or highly stylized CAs (Azevedo et al., 2018).

Figure 1. Conversational Agent (CA)

The presented medication information (e.g., purpose, dose) was grouped and ordered according to patient preferences (Morrow et al., 1996). The CA paused after presenting each category of information. In the TB condition, the CA asked a question about key concepts after each category. These questions were framed as checking CA ability to communicate rather than to test participants (to avoid the perception of being patronized) and were open-ended so that answers would better reflect whether the participant understood the CA. In the NTB condition, the same instructions were presented without the questions. After each set of instructions, memory for the information was measured using both free and cued recall procedures.

Next, participants were interviewed about the CA with scaled and open-ended questions. The scaled questions included the API, commonly used to measure responses to CAs (Ryu & Baylor, 2005). Factor analyses have identified two subscales: Informational Usefulness, or teaching effectiveness of the CA (e.g., 'agent is knowledgeable'; 'agent made my think deeply'; 'agent is tuned to my needs') and Affective Interaction with the CA (e.g., 'agent has a personality'; 'agent is expressive'). Other scaled questions about participant perceptions of the CA were developed for the study.

Participants were also asked open-ended questions about their interaction with our CA, depending on which condition they were in. Participants in both conditions were asked their first impressions of the CA, their response to the CA's appearance and behavior (e.g., gestures, head nods), the ease of understanding the CA medication instructions, and how helpful they thought the CA was. Participants in the teachback condition were also asked about their reaction to receiving teachback questions from the CA as well as the frequency of the questions during the instruction.

Answers to the interview questions were transcribed and segmented into separate idea units or comments. Two of the investigators reviewed and discussed the comments in order to develop a set of topic codes, drawing on existing taxonomies in the CA literature (Azevedo et al., 2019; Ryu & Baylor, 2005; Hayes-Roth, Maldonado & Moraes, 2002; Heidig & Clarebout, 2011). Each comment was coded separately by the two investigators for topic and valence (negative, neutral, positive) with respect to the agent (inter-rater reliability = 81.6%). Coding disagreements were resolved by consensus between the two coders and another investigator. The resulting categories and examples are presented in Table 1.

Table 1. Exemplars of comments per category and valence.

	Category	**Negative Valence**	**Positive Valence**
CA characteristics: General	Identity	"Higher than 1% [*sic*] of population (diabetes) is African American, so have two versions." "You could have something funny like a dog or a cat, I would prefer that."	"She looks like a regular nurse, assistant, or medical personal."
	Realistic	"Make her a little bit more realistic" "Computer generated. Doesn't look like a real person"	"She seems fine. She looks like a regular human." "I think it was very cool that she was like... human-like"
	Personality	"No warmth, emotion or empathy" "Appear a little friendlier"	"She was fine. She looks knowledgeable" "She is a pleasant person. "
CA characteristics: Appearance	Visual	"The spine at the background looks awfully curved with problems." [*Comment about the provider office scene behind the CA*] "She looked fairly normal; a bit too skinny for my liking". "I did not like her pants"	"Neat and well-dressed" "It wasn't distracting at all. She could be considered "professional looking" "The medical setting was nice"
	Facial Expression	"Eye blinking was distracting." "Have more facial expressions."	"She has a good eye movement." "I like the blinking."
	Gestures	"Creepy. Hand gestures are distracting." "hand gesture doesn't make sense, seems exaggerated."	"Gestures were good." "Her gestures are fine."
CA communication: Interactivity	Usefulness	"I wasn't that impressed with her"	"It's a good idea" "Pretty good to have someone to

		"A person standing here is not more helpful than a recording or hearing it."	explain info about medication when the doctors and nurses typically don't"
	Interactive	"I wanted to ask questions. So, I can clarify what she [*the agent*] is asking first. Not being asked immediately for response" "For me, I need to hear more than once. I have a short attention span."	"I got to repeat what she was telling me" "She did a teach back. She makes sure you understand each point"
	Memory	"A lot of information at one time. I am not learning through audio well. I usually read." "She gave a lot of instructions. If she could go over or write it down a couple of times would be great".	"Good way to reinforce the information." "Instead of memorizing a bunch of info and list all at once."
CA communication: Message	Language	"For the most part liked it. It was annoying that she said thank you 6 times" "Hard to remember the name of the medication."	"She was speaking standard English without an accent in words that are commonly understood." "She spoke layman's terms which you can understand."
	Content Organization	"(...) but she did not offer any info on side effects" "Not mentioned the A1C tests. Maybe also talk about your test results. "	"Did a good job of telling the basics of the medication." "Seems like she really wants you to understand why it is important, the benefits and disadvantages might be"
	Speech Delivery	"Too long a pause between questions." "Sometimes she is a little fast."	"She kept things short, brief and concise" "Pace was good. Not too slow and not too fast."
	Multimedia/ Visual Aid	"It would be good to have a visual display, a white board." "She could have samples of the medication and show it how to take."	[*There were no positive comments in this category*]

Results and Discussion

Consistent with the first prediction, participants in both groups liked the CA's appearance (1= did not like at all; 5= liked very much; 4.3 > 3.0, Wilcoxon-Mann-Whitney U-Test, p < .001) and behavior (3.8 > 3.0, p <.001). They thought the CA-presented instructions were easy to understand (1= very difficult; 5= very easy; mean= 4.3, > 3.0, Wilcoxon-Mann-Whitney U-Test, p < .001). A nonsignificant trend suggested participants thought the CA was an effective teacher (API composite score mean=4.1, > 3.0, p<.10). They were neutral with respect to the CA affective interactive properties (API composite = 3.3, > 3.0, p > .10).

Most important and consistent with the second prediction, teachback influenced older adults' responses to the CA. Participants thought the CA was a more effective teacher in the TB condition (4.4 versus 3.8; Wilcoxon-Mann-Whitney U-Test, p <.05, Cohen's d =.87), but the groups did not differ in how they perceived the CA's affective interactive properties (3.4 versus 3.3, p >.10, Cohen's d =.17). There was a nonsignificant trend in favor of the TB group for the 'ease of understanding' measure (4.4 versus 4.1; p > .10). Participants in the TB group liked the questions (mean=4.1 > 3.0, p<.001) and thought they were helpful (mean=4.5 > 3.0, p<.001). Most (83%) thought the number of questions were appropriate.

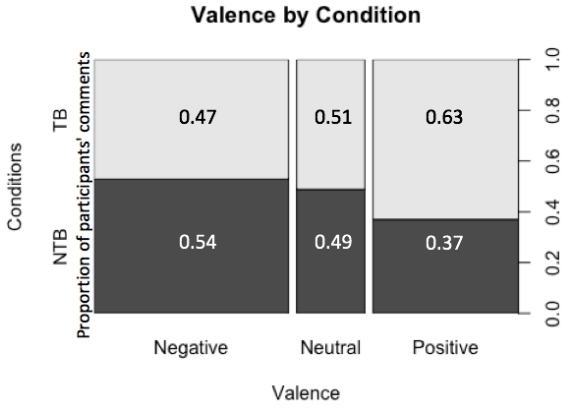

Figure 2. Overall valence by condition

Participants' comments about the CA in response to the interview questions converged with and elaborated these findings. Comments were more positive in the TB group, with a trend for more negative comments in the NTB group (Cochran Armitage Trend Test, Z=-2.29, p = 0.02, see Figure 2). Examining the trends by category (Table 2) revealed significant effects with the same pattern as for overall comments for the Interactive (Z test p=.03; Monte Carlo Correction, 10000 resamples) and for the Memory (Z, p=.01) categories. These comments suggested that participants thought the interactive CA (i.e., TB condition) helped them learn and remember the medication instructions, in part by reducing cognitive load and reinforcing key concepts.

Table 2. Frequency counts per category and valence

CA characteristics:

General: Categories	Cond.	Negative	Neutral	Positive	*p-value (adjusted)*
Identity	TB	3	0	1	
	NTB	1	3	0	1
Realistic	TB	8	1	3	
	NTB	5	1	3	0.81
Personality	TB	2	1	5	
	NTB	4	0	3	0.45
Appearance:					
Visual	TB	2	2	8	
	NTB	5	2	5	0.26
Facial Expression	TB	9	1	3	
	NTB	6	1	0	0.36
Gestures	TB	6	2	4	
	NTB	6	3	2	0.63

CA communication:

Interactivity: Categories	Cond.	Negative	Neutral	Positive	*p-value (adjusted)*
Usefulness	TB	2	1	5	
	NTB	1	0	4	0.80
Interactive	TB	4	4	6	
	NTB	7	0	1	0.03 *
Memory	TB	0	0	9	
	NTB	3	1	3	0.02 *
Message:					
Language	TB	7	1	1	
	NTB	7	2	2	0.58
Content Organization	TB	3	0	3	
	NTB	2	0	2	1
Speech Delivery	TB	3	2	8	
	NTB	6	1	9	0.69
Multimedia/ Visual Aid	TB	6	2	0	
	NTB	12	0	0	0.15

In addition to these trends favoring the CA in the TB condition, there were trends reflecting participants' overall responses to the CA. Positive comments for the Usefulness and Speech Delivery categories reflected the view that the instructions were presented clearly and at a good pace, and more generally that the CA system would help them understand how to take their medications. On the other hand, negative comments for Language, Facial Expression, Multimedia (and Gesture, to some extent) reflected the view that the CA's facial expressions and gestures were robotic and not very human-like; the medication names were difficult to remember and could be confusing; and need for multimedia/ visual aid support to elaborate the medication instruction.

Finally, we found no support for the third prediction that CA teachback would improve memory for the instructions (mean free recall: TB= 50.7% correctly recalled; NTB=54.3%; mean cued-recall: TB=73.4%; NTB=74.6%; $p>.10$ for all comparisons), possibly because teachback did not include the CA correcting participants during instruction in our prototype system.

Discussion

Our findings suggest older adults are open to interacting with CAs and that CAs could help support self-care if well designed. Participants in both groups felt the agent was generally personable and useful, although they thought the CA appearance and behavior was not sufficiently human-like. Most important, older adults thought that CA-based teachback was helpful and that the CA was a more effective teacher when using this interactive strategy, replicating findings for patients interacting with human providers who use teachback (Samuels-Kalow et al., 2016). Participants in the TB condition made more positive comments about the CA, reflecting a focus on benefits of teachback for supporting learning and remembering the medication instructions.

Thus, CAs may provide an important resource for reinforcing and augmenting provider communication about self-care and other topics, which should promote continuity of care. Next steps include developing CAs that more effectively engage patients. This includes investigating the relative impact of CA appearance, behavior, and speech for engaging older adults with their self-care through specific communication strategies such as teachback.

Acknowledgement

Research reported in this publication was approved by the Institutional Review Board at the University of Illinois at Urbana-Champaign and supported by the Jump Applied Research for Community Health through Engineering and Simulation (ARCHES) program, UIUC/OSF Hospital, Peoria, IL. The content is solely the responsibility of the authors and does not necessarily reflect the official views of these institutions.

References

Agency for Healthcare Research and Quality (AHRQ) (2019, March). More effort is needed to ensure patients understand doctors' instructions. Retrieved February 23, 2020 from https://www.ahrq.gov/sites/default/files/wysiwyg/research/findings/nhqrdr/dataspotlight-health-literacy.pdf

Azevedo, R.F.L., Morrow, D., Graumlich, J., Willemsen-Dunlap, A., Hasegawa-Johnson, M., Huang, T.S., … Halpin, D.J. (2018). Using computer agents to explain medication instructions to older adults. *AMIA Annual Symposium Proceedings.* Washington, DC: American Medical Informatics.

Azevedo, R. F. L., Morrow, D., Gu, K., Huang, T., Hasegawa-Johnson, M., Soni, P., Tang, S., Sakakini, T., Bhat, S., Willemsen-Dunlap, A., & Graumlich, J. (2019). The Influence of Computer Agent Characteristics on User Preferences in Health Contexts. *Proceedings of the 2019 Human Factors and Ergonomics Society Health Care Symposium*. Chicago, IL: The Human Factors and Ergonomics Society.

Bickmore, T. W., Pfeifer, L. M., Byron, D., Forsythe, S., Henault, L. E., Jack, B. W., ... & Paasche-Orlow, M. K. (2010). Usability of conversational agents by patients with inadequate health literacy: evidence from two clinical trials. *Journal of Health Communication*, *15*(S2), 197-210.

Chew, L. D., Bradley, K. A., & Boyko, E. J. (2004). Brief questions to identify patients with inadequate health literacy. *Family Medicine*, *36(8)*, 588-594.

Hartholt, A., Traum, D., Marsella, S., Shapiro, A., Stratou, G., Leuski, A., Morency, L-P., & Gratch, J. (2013). All together now: introducing the Virtual Human Toolkit. *Proceedings of the 2013 International Conference on Intelligent Virtual Humans*.

Hayes-Roth, B., Maldonado, H., & Moraes, M. C. (2002). Designing for diversity: Multi-cultural characters for a multi-cultural world. In: *the European Digital Content Creation Event. IMAGINA' 02* (pp. 207-225), France.

Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? *Education Research Review, 6*, 27-54.

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, *331*(6018), 772-775.

Medlock, S., Eslami, S., Askari, M., Arts, D. L., Sent, D., de Rooij, S. E., & Abu-Hanna, A. (2015). Health information–seeking behavior of seniors who use the internet: a survey. *Journal of Medical Internet Research*, *17*, e10.

Morrow, D. G., Leirer, V. O., Andrassy, J. M., Tanke, E. D., & Stine-Morrow, E. A. (1996). Medication instruction design: Younger and older adult schemas for taking medication. *Human factors*, *38*(4), 556-573.

Paasche-Orlow, M. K., Schillinger, D., Greene, S. M., & Wagner, E. H. (2006). How health care systems can begin to address the challenge of limited literacy. *Journal of General Internal Medicine*, *21*(8), 884-887.

Ryu, J., & Baylor, A. L. (2005). The psychometric structure of pedagogical agent persona. *Technology Instruction Cognition and Learning*, *2*(4), 291.

Samuels-Kalow, M., Hardy, E., Rhodes, K., & Mollen, C. (2016). “Like a dialogue”: teach-back in the emergency department. *Patient Education and Counseling*, *99*(4), 549-554.

Schillinger, D., Piette, J., Grumbach, K., Wang, F., Wilson, C., Daher, C., ... & Bindman, A. B. (2003). Closing the loop: physician communication with diabetic patients who have low health literacy. *Archives of internal medicine, 163*(1), 83-90.

Straßmann, C., & Krämer, N. C. (2017, August). A categorization of virtual agent appearances and a qualitative study on age-related user preferences. In *International Conference on Intelligent Virtual Agents* (pp. 413-422). Springer, Cham.

Street Jr, R. L., Makoul, G., Arora, N. K., & Epstein, R. M. (2009). How does communication heal? Pathways linking clinician–patient communication to health outcomes. *Patient Education and Counseling*, *74*(3), 295-301.

Wagner, E. H., Austin, B. T., Davis, C., Hindmarsh, M., Schaefer, J., & Bonomi, A. (2001). Improving chronic illness care: translating evidence into action. *Health Affairs, 20*(6), 64-78.

White, M., Garbez, R., Carroll, M., Brinker, E., & Howie-Esquivel, J. (2013). Is “teach-back” associated with knowledge retention and hospital readmission in hospitalized heart failure patients? *Journal of Cardiovascular Nursing*, *28*(2), 137-146.