

# Health Document Explanation by Virtual Agents

Timothy W. Bickmore<sup>1</sup>, Laura M. Pfeifer<sup>1</sup> and Michael K. Paasche-Orlow, MD<sup>2</sup>

<sup>1</sup>Northeastern University College of Computer and Information Science  
360 Huntington Ave WVH202, Boston, MA 02115  
(bickmore, laurap)@ccs.neu.edu  
<http://www.ccs.neu.edu/research/rag/>

<sup>2</sup>Boston University School of Medicine  
91 East Concord St, Suite 200, Boston, MA 02118  
mpo@bu.edu

**Abstract.** We describe the design and evaluation of a virtual agent that explains health documents to patients. The prevalence and impact of low health literacy is presented as a motivation for such agents, given that face-to-face interaction with health providers is cited as one of the most effective means of communicating with these patients. We analyze the form and distribution of pointing gestures used by experts in explaining health documents, and use this data to develop a computational model of agent-based document explanation. This model is evaluated in a randomized controlled trial. Preliminary results indicate that patients with low health literacy are more satisfied with health document explanation by a virtual agent compared to a human.

**Keywords:** virtual agent, relational agent, embodied conversational agent, health literacy, hand gesture.

## 1 Introduction

Many professionals provide their lay clients with documents that are, to varying degrees, incomprehensible. Whether due to technical jargon, obscure concepts, or poor writing on the part of the professional, or low literacy, cultural barriers, or cognitive impairment on the part of the client, documents often fail to serve their intended communicative function.

Perhaps nowhere is this problem more important and pervasive than in healthcare. The consequences of a patient failing to understand a prescription, hospital discharge instructions, or pre-surgery instructions can have serious, even fatal, consequences. The inappropriate complexity of documents has been discussed in the medical literature for over 50 years and in the past two decades this has been broadly recognized as a serious problem within the US medical community [5]. Indeed, a significant and growing body of research has emerged relating to the problem of “health literacy”, which has brought attention to the ethical and health impact of overly complex documents in healthcare [30]. Virtual agents may provide a particularly effective solution for addressing this problem, by having the agents describe health documents to patients using exemplary techniques that an expert health provider might use, given that

they had training in communicating with patients with low health literacy and had unbounded time available.

In this paper we describe our initial efforts in building and evaluating a virtual agent designed to explain health documents to patients.

### **1.1 Health Literacy**

Health literacy is the ability to perform the basic reading and numerical tasks required to function in the health care environment, and it affects patients' ability to understand medication labels and instructions, hospital discharge instructions, instructions for assistive devices and medical equipment, and health education material [1]. Patients with inadequate health literacy report lower health status [40], are less likely to use screening procedures, follow medical regimens, keep appointments, or seek help early in the course of a disease [39], have greater difficulties naming their medications and describing their indications [4,41], more frequently hold health beliefs that interfere with adherence [18], have higher health-care costs [39], and have higher rates of hospitalization [4]. Fully 90 million American adults have limited literacy skills and limited literacy has been shown to be more prevalent among patients with chronic diseases, those who are older, minorities, and those who have lower levels of education [31]. Seminal reports about the problem of health literacy include a sharp critique of current norms for overly complex documents in health care such as informed consent [1,28].

### **1.2 Virtual Agents to Address Low Health Literacy**

Evidence suggests that face-to-face encounters with a health provider—in conjunction with written instructions—remains one of the best methods for communicating health information to patients in general, but especially those with low literacy levels [11,23,25,33]. Face-to-face consultation is effective because it requires that the provider focus on the most salient information to be conveyed [33] and that the information be delivered in a simple, conversational speaking style. Protocols for grounding in face-to-face conversation allow providers to dynamically assess a patient's level of understanding and repeat or elaborate information as necessary [10]. Face-to-face conversation also allows providers to make their communication more explicitly interactive by asking patients to do, write, say, or show something that demonstrates their understanding [14]. Finally, face-to-face interaction allows providers to use verbal and nonverbal behaviors, such as empathy [15] and immediacy [34], to elicit patient trust, enabling better communication and satisfaction.

Of course, one problem with in-person encounters with health professionals is that all providers function in health care environments in which they can only spend a very limited amount of time with each patient [12]. Time pressures can result in patients feeling too intimidated to ask questions, or to ask that information be repeated. Another problem is that of "fidelity": providers do not always perform in perfect accordance with recommended guidelines, resulting in significant inter-provider and intra-provider variations in the delivery of health information.

Given the efficacy of face-to-face consultation, one technology that shows particular promise for conveying health information to patients with low health literacy is the use of virtual agents that simulate face-to-face conversation with a provider. These systems can recognize and produce verbal and nonverbal conversational behaviors that signify understanding and mark significance, and can convey information in redundant channels of information (e.g., hand gestures, such as pointing, facial display of emotion, and eye gaze), to maximize message comprehension. They can use the verbal and nonverbal communicative behaviors used by providers to establish trust and rapport with their patients in order to increase satisfaction and adherence to treatment regimens [6]. They can adapt their messages to the particular needs of patients and to the immediate context of the conversation. Virtual agents can provide health information in a consistent manner and in a low-pressure environment in which patients are free to take as much time as they need to thoroughly understand it. This is particularly important as health providers frequently fail to illicit patients' questions, and patients with low literacy are even less likely than others to ask questions [19].

Virtual agents can also consistently evaluate patient comprehension of the material presented. Physicians infrequently evaluate patients' understanding, and when they do it is mostly simply to ask "do you understand?" without waiting for a reply [36].

## **2 Related Work**

Virtual pedagogical agents and virtual agents that use deictic (pointing) hand gestures to refer to objects in their virtual or physical environment represent the two areas of previous research that are most relevant to our work. An agent that explains a document is essentially teaching the user about the topics covered in the document, and thus pedagogical strategies pioneered by other developers of virtual agents are of interest. The appropriate use of deictic gestures is particularly important for document explanation, since they are required to orient the user to the part of the document under discussion.

Virtual pedagogical agents include Autotutor [16], Steve [35], Cosmo [21], Persona [3], Sam [7] and others. Evaluations of these agents have largely shown mixed educational outcomes. For example, users rated the Persona agent as more entertaining and helpful than an equivalent interface without the agent [3]. However, there was no difference in actual performance (comprehension and recall of presented material) in interfaces with the agent vs. interfaces without it. On the other hand, researchers evaluating the Cosmo agent found that: 1) students who interacted with an educational software system with a pedagogical agent produced more correct solutions and rated their motivation to continue learning and interest in the material significantly higher, compared to the same system without the agent; 2) students who interacted with an agent that used speech output, rated the lessons more favorably and recalled more compared with students who interacted with an agent that used text output; and 3) students who interacted with an agent that used personalized dialogue recalled more than students who interacted with an agent that communicated using non-personalized monologues (as in video-based education) [24]. In another study, stu-

dents using the AutoTutor pedagogical agent in addition to their normal coursework outperformed both a control group (no additional intervention), and a group directed to re-read relevant material from their textbooks [32].

Deictic gestures represent perhaps the most common type of hand gesture implemented in virtual agents. Early examples include the Persona agent, which could point to parts of images on web pages, and Jack, the virtual meteorologist, who could point at weather images that he stood in front of (in his virtual environment) while giving a weather report [29]. However, the interaction and gesture specifications in these early systems were mostly scripted. The BEAT system incorporated a simple rule that generated deictics whenever a new object in the agent's virtual world was referenced in speech and was "visible" to both the agent and user [9]. Perhaps the most sophisticated model of "deictic believability" was implemented for the Cosmo agent [22]. This system utilized a separate deictic planner that would determine the generation of deictics on the basis of speech act, gesture referent, speech referent, world model (including possible distractors) and discourse history. Virtual agents that can use alternative modalities have also used deictic gestures: MACK could highlight a paper map that was "shared" with a user by means of an overhead projector [8], and Steve accompanied users into a virtual reality world where he could point out virtual objects that the user needed to manipulate [35].

To our knowledge, with the exception of the maps used by MACK, no virtual agent has been designed to date for the task of explaining physical documents to users.

### 3 Health Document Explanation by Human Experts

In order to develop a virtual agent that can emulate expert document explanation behavior, we analyzed four example interactions in which experts were explaining health documents to others (Fig. 1). Two of these conversations involved a hospital discharge procedure, in which a nurse reviewed diagnoses, medications, follow-up appointments, and self-care procedures with a patient by explaining an "After Hospital Care Plan" (AHCP) document. The other two conversations involved a research assistant explaining a research informed consent (CONSENT) document. The



**Fig 1.** Explanation of AHCP (left) and CONSENT (right) by experts.

CONSENT document was two pages long and consisted entirely of text, mostly in non-technical language, whereas the AHCP was eleven pages long and consisted of a mixture of text and images. While the AHCP was explicitly designed for patients with low health literacy, it is full of medical terminology (medication names, medical condition names, etc.). We created a single standard instance of AHCP and CONSENT documents that were used in these studies and all subsequent evaluations. All four interactions were “mock” conversations in that the listener was another research assistant, but the four individuals doing the explaining were experts in their respective areas. In the two AHCP examples, the nurse and “patient” are seated next to each other at a table with the document on the table between them. In the CONSENT examples, the research assistant and “patient” are seated facing each other, and the research assistant holds the document up for the patient. All four interactions were videotaped for subsequent analysis.

The videos were transcribed and broken into utterances, following [27]. Speech acts were coded for each utterance using the DAMSL coding scheme [2]. Table 1 provides an overview of the four conversations.

Conversation	Document	TRT	Utterances		
			Expert	Patient	Total
1	CONSENT	2:08	93	1	94
2	CONSENT	2:24	103	8	111
3	AHCP	6:46	282	32	314
4	AHCP	6:53	277	39	316

**Table 1.** Conversations Analyzed

### 3.1 Analysis of Nonverbal Behavior During Human Document Explanation

Given the importance of face-to-face interaction in communicating with low literacy patients, we focused our initial analysis and modeling efforts on the nonverbal behavior of the expert in these conversations. Initial reviews of the videos indicated that one expert behavior was ubiquitous: deictic (pointing) gestures referencing the document. Thus, we further focused our initial analysis on the occasioning and form of these deictic gestures.

The start and end of each expert gesture was coded from the video, along with the form of gesture. The timing of gesture stroke relative to utterance was also coded as: before utterance, beginning of utterance (first three words), ending of utterance (last 3 words), middle of utterance, or continued from previous utterance (following [26]). We observed a wide range of deictic forms, but decided to initially collapse these into POINT (pointing and underlining) and REGION (whole hand) gestures.

Preliminary analyses indicated that a verbal reference to a new part of the document (relative to the one currently under discussion) seemed to be a good predictor for deictic gesture. Consequently, we also coded the part of the document under discussion by the expert. The documents were broken up into topic level by identifying pages, regions and items within each document. Each topic was represented by an ID number in the format “<page>.<section>.<item>”, eg. “1.4.2”. We also created a

code to indicate the topic level being introduced (PAGE, SECTION, or ITEM), as well as a code that indicated relative navigation in the document (IN, OUT, FORWARD, etc.), both based on changes in the topic ID,

Chi-squared tests for independence indicated that speech act, topic level, and document navigation were all strongly associated with the occurrence and form of deictic gesture performed during a given utterance (NONE, POINT or REGION,  $p < .001$ ). We then used a commercial decision tree modeling tool (DTREG.com) to evaluate models based on various combinations of these coded predictors. The lowest error rate found (15.5%) was for a model that considered all available information (speech act, topic level, etc.). However, the model based on topic level alone was only slightly worse (15.6% error rate), so we decided to base our initial computational model on topic level alone to simplify implementation.

### 3.2 Resulting Model of Deictic Gesture during Document Explanation

Our preliminary model generates a deictic gesture according to the model in Table 2.

New Topic Level	Gesture		
	NONE	POINT	REGION
<b>No Change</b>	92.8%	4.4%	2.8%
<b>PAGE</b>	57.7%	3.8%	38.5%
<b>SECTION</b>	23.7%	36.8%	39.5%
<b>ITEM</b>	23.7%	21.1%	55.3%

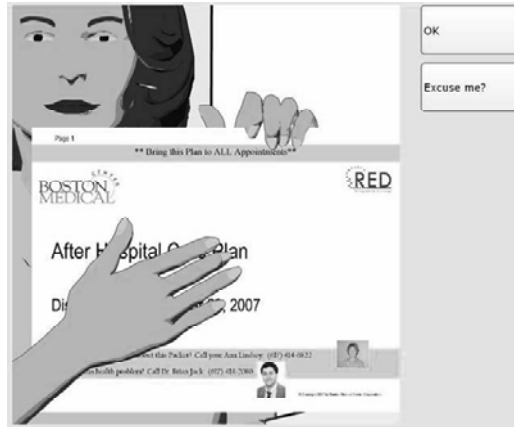
**Table 2.** Document Deictic Generation Model

We found that 83% of the time, deictic gesture stroke occurred at the beginning of an utterance. Thus, in our model, all document deictic gestures are generated with their stroke at the beginning of the related utterance.

## 4 A Computational Model of Document Explanation

An existing virtual agent framework designed for health counseling [6] was modified to provide explanation of health documents. The framework features a vector-graphics-based virtual agent whose nonverbal behavior is synchronized with a text-to-speech engine (Fig. 2). User contributions to the conversation are made via a touch screen selection from a multiple choice menu of utterance options, updated at each turn of the conversation.

Dialogues are scripted, using a custom hierarchical transition network-based scripting language based on Augmented Transition Networks (ATNs) [42]. ATNs are used both to model the hierarchical structure of dialogue [17] as well as to enable common sub-dialogues to be factored out and re-used. In addition to network branching operations, ATN actions can include saving values to a persistent database or retrieving and testing values from the database, in order to support the ability to remember and refer back to information from earlier turns and prior conversations. Agent utterances can



**Fig 2. Virtual Agent Interface**



**Fig 3. Hospital Cart**

be tailored at runtime through the inclusion of phrases derived from information in the database or other sources (template-based text generation).

The virtual agent has a range of nonverbal behaviors that it can use, including: hand gestures, body posture shifts, gazing at and away from the user, raising and lowering eyebrows, head nods, different facial expressions, and variable proximity (wide to close-up camera shots). Co-verbal behavior is determined for each utterance using the BEAT text-to-embodied-speech system [9], with several enhancements to support health dialogues. One such enhancement is that conversational frame [38] (task-oriented, social, empathetic, or encouraging) can be specified in the script and automatically translated into appropriate changes to facial expression, proximity, speech synthesizer intonation, and gesture frequency output by BEAT. While we are aware of some of the limitations of BEAT [20], we find that it is adequate for our purposes (e.g., we have yet to encounter a need for partially-overlapping temporal spans of verbal and nonverbal behavior).

The framework was extended for document explanation in several ways. Two new characters (“Elizabeth” and “Louise”) were developed for this application that were more concordant with the hospital population we intend to deploy it in (middle aged, female, Caucasian and African American). A set of animation system commands was added to allow document pages to be displayed by the character (Fig. 1), with page changes automatically accompanied by a page-turning sound. A set of document deictic gestures was added so that the agent could be commanded to point anywhere in the document with either a pointing hand or an open hand. While the document is displayed, the agent can continue using its full range of head and facial behavior, with gaze-aways modified so that the agent looks at the document when not looking at the user (in our corpus, the expert gazed at the document 65% of the time and at the patient 30% of the time). However, hand gestures were limited to document deictics, and posture and proxemic shifts were disabled while the document is displayed.

In the first version of the system, we were primarily interested in the generation of document deictics given the verbal content of the document explanation script. In order to provide the information needed by BEAT for this, we introduced the use of an

additional XML tag in every agent utterance that referenced the document. These tags specified the document location ID described in Section 3.1, and the X,Y coordinates (normalized to 100%, 100%) of the page corresponding to the location ID, for example:

```
<DOC PART="2.1.1" LOC="25,40"> It is for your blood  
pressure. </DOC>
```

A BEAT behavior generator was created that tracked document context (current and previous document locations under discussion) and annotated the utterance parse tree with: page change specifications (whenever the document location ID indicated a change in page); document deictic gestures (per the rules described in Section 3.2); and additional gaze-aways (at the start of all utterances in which a document deictic gesture or page change is indicated).

## 5 Preliminary Evaluation Study

We conducted a pilot evaluation study to test the efficacy of our agent-based document explanation system, compared with a standard of care control (explanation by a human) and a non-intervention control (self study of the document in question). The study had a 3 (AGENT vs. HUMAN vs. SELF) x 2 (AHCP vs. CONSENT) between-subjects experimental design, in which each participant evaluated two different conditions in a single session, always AHCP followed by CONSENT, with the presentation of the other conditions randomized. This document ordering was intended to minimize carryover effects from the informed consent procedure for the pilot study itself to the CONSENT treatment of the study. To further minimize carryover, we used an informed consent document from an entirely different area of medical research (acquisition of blood samples for genetic banking). The study was approved by Northeastern University's IRB.

### 5.1 Apparatus

Two interaction scripts were created, one for the AHCP and one for CONSENT, based on the videotapes described in Section 3. In each script, users could simply advance linearly through the explanation (by selecting "OK"), ask for any utterance to be repeated ("Could you repeat that please?"), request major sections of the explanation to be repeated, or request that the entire explanation be repeated. Any number of repeats could be requested and, although the scripting language has the ability to encode rephrasings when an utterance is repeated, for the current study the agent would repeat the exact same utterance when a repeat was requested for any state in the script. The agent was deployed on a mobile cart with a touch screen attached via an articulated arm (Fig. 2), since this is the platform we will be using in the hospital for pre-discharge patient education (the articulated arm enables the screen to be positioned in front of a patient in a hospital bed). Study sessions were held in an observation room of our HCI laboratory, with the interactions videotaped using four closed-circuit video cameras.



## 5.2 Measures

In addition to basic demographics, we assessed health literacy using the REALM instrument, which categorizes individuals into 3<sup>rd</sup> grade and below, 4<sup>th</sup>-6<sup>th</sup> grade, 7<sup>th</sup>-8<sup>th</sup> grade, and high school [13]. We also created knowledge tests for each of the two documents, with the one for CONSENT based on the BICEP evaluation [37]. Note that these tests were always administered in an “open book” fashion with the participant able to refer to a paper copy of the document during the test. We augmented the BICEP with scale measures of likelihood to sign the consent document and perceived pressure to sign the consent document.

Evaluation questionnaires were also developed for the HUMAN and AGENT study conditions, assessing satisfaction with the instructor and with the overall instructional experience, desire to continue working with the instructor, trust in the instructor, and how knowledgeable the instructor was, all evaluated on 7-point scales.

## 5.3 Participants

Eighteen subjects participated in the study, were recruited via fliers posted around the Northeastern University campus, and were compensated for their time. Participants had to be 18 years of age or older and able to speak English. Participants were 74% male, aged 19-33. Two were categorized as 4<sup>th</sup>-6<sup>th</sup> grade, three as 7<sup>th</sup>-8<sup>th</sup> grade, and the rest as high school level, according to the REALM health literacy instrument.

## 5.4 Procedure

Participants arrived at the HCI laboratory, were consented, filled out the demographic questionnaire and then had the REALM health literacy evaluation administered.

Following this they were exposed to one of the three experimental conditions for the AHCP document. For the AGENT condition, they were given a brief training session on how to interact with the agent, the experimenter then gave the participant a paper copy of the document, left the room and closed the door. At the end of the interaction the virtual agent informed the participant that they could take as much time as they liked to review the document before signaling to the experimenter that they were ready to continue. For the HUMAN condition, a second research assistant in our lab explained the document to the study participant. This instructor did not have a health care background, but routinely administered informed consent for HCI studies and was allowed to watch the videotapes described in Section 3 to learn about the AHCP. The instructor was blind to the virtual agent interaction script content and evaluation instruments, and was simply asked to explain the document in question to the participant. For the CONTROL condition, the participant was simply handed the document and told to take as much time as they needed to read and understand it, and were then left alone in the observation room until they signaled they were ready to continue.

Following the first intervention, the research assistant verbally administered the AHCP knowledge test and instructor evaluations. The previous two steps were then repeated with the CONSENT document.

### 5.5 Results

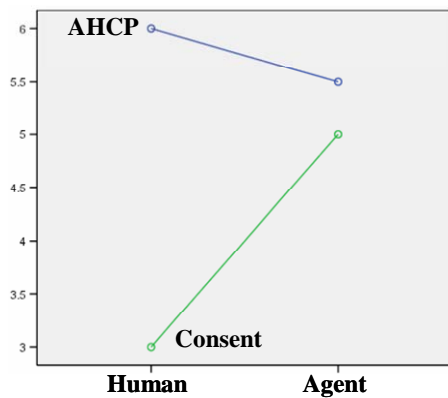
We conducted full-factorial ANOVAs for all measures, with condition (AGENT, HUMAN, SELF), document (AHCP, CONSENT) and health literacy (four categories) as independent factors, and LSD post-hoc tests when applicable.

There was one main effect of document on test score (66.3 vs. 87.1,  $F(1,18)=14.5, p<.001$ ) indicating that participants scored significantly higher on the AHCP test compared to the CONSENT test. There were no significant effects of condition or literacy on test score.

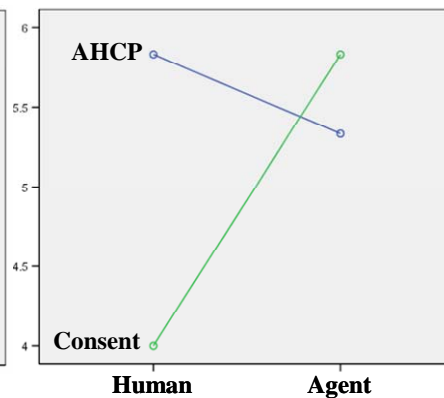
Instructor evaluations for the AGENT and HUMAN conditions indicated a number of significant effects. There was a significant interaction between condition, document and literacy on satisfaction with the overall experience ( $F(1,14)=5.0, p<.05$ ) such that those in the highest literacy level were more satisfied with the agent compared to the human for CONSENT, but were more satisfied with the human for AHCP (Fig 4). However, lower literacy participants were more satisfied with the agent in all situations (5.14 for HUMAN vs. 6.33 for AGENT).

All participants rated the agent as more knowledgeable than the human for CONSENT, but the human more knowledgeable for AHCP (Fig 5,  $F(1,140)=6.0, p<.05$ ).

There were also several main effects for literacy, with the lowest literacy participants scoring significantly lower on trust in the instructor (whether human or agent,  $F(2,14)=4.4, p<.05$ ), how knowledgeable the instructor was ( $F(2,14)=3.8, p<.05$ ), and desire to continue working with the instructor ( $F(2,14)=4.2, p<.05$ ). For the CONSENT document, there were significant effects of health literacy on likelihood to sign ( $F(2,12)=6.4, p<.05$ ) and perceived pressure to sign ( $F(2,12)=132.0, p<.001$ ), such that those with lowest literacy were significantly less likely to sign and felt sig-



**Fig 4.** Satisfaction with Overall Experience by Participants with High School Level Literacy



**Fig 5.** Knowledge of Instructor Ratings by All Participants

nificantly more pressure to sign, compared to those with higher levels of literacy.

Six participants interacted with both the human and the agent in a single session, so we also compared ratings from these participants using a matched-pair analysis for increased power. These participants rated the agent significantly higher on satisfaction with the instructor (6.0 vs. 5.17, paired  $t(5)=2.7$ ,  $p<.05$ ) and satisfaction with the overall experience (6.17 vs. 5.0, paired  $t(5)=2.9$ ,  $p<.05$ ), compared to the human.

## 6 Conclusions

Although we did not see significant differences in test scores across intervention conditions, this was not too surprising given the relatively high literacy status of the participants and the fact that the tests were “open book”. However, in a busy clinic, especially with low literacy patients, the agent may actually outperform a time constrained and impatient clinician. As some participants put it:

- “I’d rather have Elizabeth. I liked the interface. I liked the way the tone has been set to explain to people. It doesn’t kind of exert too much pressure on the person who’s listening, so I like that.”
- “Elizabeth was cool, I would have taken that again. She was just so clear, she just went page by page so it wasn’t missed. And then, I mean you can always just ask them [human] if you don’t understand anyway, but it’s different on a screen, I guess, because some people don’t want to say that they don’t understand. On a screen it’s less embarrassing, no one’s here so you can say ‘Ok, let me hear that again.’”
- “Honestly I would have rather had Elizabeth explain them to me, just because of uh, maybe a situation with a conflict of interest, or maybe the distractions or something like that. Because everybody has distractions that keep us astray from the concept underlying the document.”

While we did not see effects on test scores, we did see clear patterns emerge on satisfaction, with direct comparisons by participants who interacted with both the agent and human, as well as all evaluations by low literacy participants, indicating a preference for the agent.

The higher scores for the human describing the AHCP (satisfaction for high literacy participants and ratings of how knowledgeable the instructor was for all participants) may be due to the increased length and complexity of the AHCP and its much higher density of medical terms:

- Researcher: “Did you feel that Elizabeth did a better job explaining one of the documents to you versus another?” Participant: “Probably, yes, the second one. Well, maybe because it was easier to understand because it had less terminology.”

One interpretation of these results is that the script for agent-based AHCP explanation was inadequate along one or more dimensions, and that only participants with high literacy were able to notice the deficiencies. A more likely explanation (referring to Figs. 4 and 5) is that participants did not like the human research assistant’s explanation of the consent document.

In sum, we feel that the pilot study indicates that document explanation may be a very important application domain for virtual agents, and health document explanation to patients with low health literacy may be particularly significant.

Our future work is focused on extending the study with participants in the lowest literacy category. We also plan to revise the computational model using data from real provider-patient interactions (especially patients with low literacy), expand the repertoire and precision of the models of nonverbal behavior, and begin to investigate the automatic generation of explanation dialogue given a document as input. Finally, the preliminary results from this first study must be further validated through additional testing with a broader range of explanation scripts and a larger sample of human experts performing document explanation for comparison.

**Acknowledgements.** Thanks to Francisco Crespo and Thomas Brown for their assistance in conducting the evaluation study, and to our collaborators at Boston Medical Center—Dr. Brian Jack and Anna Johnson—for providing the example After Hospital Care Plan and videotaping the mock hospital discharge consultations. Jennifer Smith provided many helpful comments on the paper. This work was supported by a grant from the NIH National Heart Lung and Blood Institute.

## References

1. Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, A. M. A.: Health literacy: report of the Council on Scientific Affairs. *JAMA* 281, 6 (1999) 552-557.
2. Allen, J. and Core, M.: Draft of DMSL: Dialogue Act Markup in Several Layers. (1997).
3. Andre, E., Rist, T., and Muller, J.: Integrating reactive and scripted behaviors in a life-like presentation agent. *Proceedings of AGENTS'98* (1998) 261-268.
4. Baker, D., Parker, R., Williams, M., Clark, S., and Nurss, J.: The relationship of patient reading ability to self-reported health and use of services. *Am J Public Health* 87, (1997) 1027-1030.
5. Beecher, H. K.: Ethics and clinical research. *N.Engl.J.Med.* 274, 24 (1966) 1354-1360.
6. Bickmore, T. and Picard, R.: Establishing and Maintaining Long-Term Human-Computer Relationships. *ACM Transactions on Computer Human Interaction* 12, 2 (2005) 293-327.
7. Cassell, J., Ananny, M., Basu, A., Bickmore, T., Chong, P., Mellis, D., Ryokai, K., Smith, J., Vilhjálmsón, H., and Yan, H.: Shared Reality: Physical Collaboration with a Virtual Peer. *Proceedings of CHI '00* (2000)
8. Cassell, J., Stocky, T., Bickmore, T., Gao, Y., Nakano, Y., Ryokai, K., Tversky, D., Vaucelle, C., and Vilhjálmsón, H.: MACK: Media lab Autonomous Conversational Kiosk. *Imagina '02* (2002)
9. Cassell, J., Vilhjálmsón, H., and Bickmore, T.: BEAT: The Behavior Expression Animation Toolkit. *SIGGRAPH '01* (2001) 477-486.
10. Clark, H. H. and Brennan, S. E.: Grounding in Communication. In L. B. Resnick, J. M. Levine, and S. D. Teasley, Eds.: *Perspectives on Socially Shared Cognition*, American Psychological Association, Washington (1991) 127-149.
11. Clinite, J. and Kabat, H.: Improving patient compliance. *J Am Pharm Assoc* 16, (1976) 74-76.
12. Davidoff, F.: Time. *Ann Intern Med* 127, (1997) 483-485.

13. Davis, T. C., Long, S. W., Jackson, R. H., Mayeaux, E. J., George, R. B., Murphy, P. W., and Crouch, M. A.: Rapid estimate of adult literacy in medicine: a shortened screening instrument. *Fam Med* 25, 6 (1993) 391-5.
14. Doak, C., Doak, L., and Root, J., *Teaching patients with low literacy skills*, 2nd ed., JB Lippincott, Philadelphia, PA (1996).
15. Frankel, R.: Emotion and the Physician-Patient Relationship. *Motivation and Emotion* 19, 3 (1995) 163-173.
16. Graesser, A. and al, e.: AutoTutor: A simulation of a human tutor. *Cognitive Systems Research* 1, 1999).
17. Grosz, B. and Sidner, C.: Attention, Intentions, and the Structure of Discourse. *Computational Linguistics* 12, 3 (1986) 175-204.
18. Kalichman, S., Ramachandran, B., and Catz, S.: Adherence to a combination antiretroviral therapies in HIV patients of low health literacy. *J Gen Intern Med* 14, (1999) 267-273.
19. Katz, M. G., Jacobson, T. A., Veledar, E., and Kripalani, S.: Patient Literacy and Question-asking Behavior During the Medical Encounter: A Mixed-methods Analysis. *J Gen Intern Med.* (2007).
20. Kopp, S., Krenn, B., Marsella, S., Marshall, A. N., Pelachaud, C., Pirker, H., Thórisson, K., and Vilhjálmsson, H.: Towards a Common Framework for Multimodal Generation: The Behavior Markup Language. *Intelligent Virtual Agents*, Marina Del Rey, CA (2006).
21. Lester, J., Towns, S., Callaway, C., Voerman, J., and Fitzgerald, P.: Deictic and Emotive Communication in Animated Pedagogical Agents. In J. Cassell, Ed., *Embodied Conversational Agents*, MIT Press, Cambridge, MA (2000)
22. Lester, J., Voerman, J., Towns, S., and Callaway, C.: Deictic Believability: Coordinating Gesture, Locomotion, and Speech in Lifelike Pedagogical Agents. *Applied Artificial Intelligence* 13, 4-5 (1999) 383-414.
23. Madden, E.: Evaluation of outpatient pharmacy patient counseling. *J Am Pharm Ass* 13, (1973) 437-443.
24. Moreno, R., Lester, J. C., and Mayer, R. E.: Life-Like Pedagogical Agents in Constructivist Multimedia Environments: Cognitive Consequences of their Interaction. *ED-MEDIA* (2000) 741-746.
25. Morris, L. and Halperin, J.: Effects of Written Drug Information on Patient Knowledge and Compliance: A Literature Review. *Am J Public Health* 69, 1 (1979) 47-52.
26. Nakano, Y. I., Reinstein, G., Stocky, T., and Cassell, J.: Towards a Model of Face-to-Face Grounding. Annual Meeting of the Association for Computational Linguistics (2003)
27. Nakatani, C. and Traum, D.: Coding discourse structure in dialogue (version 1.0). University of Maryland, (1999).
28. Nielsen-Bohlman, L., Panzer, A. M., Hamlin, B., and Kindig, D. A.: Institute of Medicine. *Health Literacy: A Prescription to End Confusion*. Committee on Health Literacy, Board on Neuroscience and Behavioral Health. National Academies Press, Washington DC (2004).
29. Noma, T., Zhao, L., and Badler, N. I.: Design of a virtual human presenter. *Computer Graphics and Applications*, IEEE 20, 4 (2000) 79-85.
30. Paasche-Orlow, M., Greene, S. M., and Wagner, E. H.: How health care systems can begin to address the challenge of limited literacy. *J Gen Intern Med.* 21, 8 (2006) 884-887.
31. Paasche-Orlow, M. K., Parker, R. M., Gazmararian, J. A., Nielsen-Bohlman, L. T., and Rudd, R. R.: The prevalence of limited health literacy. *J Gen Intern Med* 20, 2 (2005) 175-184.
32. Person, N. K., Graesser, A. C., Bautista, L., and Mathews, E. C.: Evaluating Student Learning Gains in Two Versions of AutoTutor. In J. D. Moore, C. L. Redfield, and W. L. Johnson, Eds.: *Artificial intelligence in education: AI-ED in the wired and wireless future*, IOS Press, Amsterdam (2001) 286-293.

33. Qualls, C., Harris, J., and Rogers, W.: Cognitive-Linguistic Aging: Considerations for Home Health Care Environments. In W. Rogers and A. Fisk, Eds.: *Human Factors Interventions for the Health Care of Older Adults*, Lawrence Erlbaum, Mahwah, NJ (2002) 47-67.
34. Richmond, V. and McCroskey, J.: Immediacy. In *Nonverbal Behavior in Interpersonal Relations*, Allyn & Bacon, Boston (1995) 195-217.
35. Rickel, J. and Johnson, W. L.: *Animated Agents for Procedural Training in Virtual Reality: Perception, Cognition and Motor Control*. Applied Artificial Intelligence (1998).
36. Schillinger, D., Piette, J., Grumbach, K., Wang, F., Wilson, C., and Daher, C.: Closing the loop: physician communication with diabetic patients who have low health literacy. *Arch.Intern.Med* 163, 1 (2003) 83-90.
37. Sugarman, J., Lavori, P. W., Boeger, M., Cain, C., Edson, R., Morrison, V., and Yeh, S. S.: Evaluating the quality of informed consent. *Clinical Trials* 2, 1 (2005) 34.
38. Tannen, D., Ed., *Framing in Discourse*. Oxford University Press, New York (1993).
39. Weiss, B.: Illiteracy among Medicaid recipients and its relation to health care costs. *J Health Care Poor Underserved* 4, 5 (1994) 99-111.
40. Weiss, B., Hart, G., McGee, D., and D'Estelle, S.: Health status of illiterate adults: relation between literacy and health status among persons with low literacy skills. *J Am Board Fam Pract* 5, (1992) 254-257.
41. Williams, M., Parker, R., Baker, D., Coates, W., and Nurss, J.: The impact of inadequate functional health literacy on patients' understanding of diagnosis, prescribed medications, and compliance. *Acad Emerg Med* 2, (1995) 386.
42. Woods, W. A.: Transition Network Grammars for Natural Language Analysis. In B. J. Grosz, K. S. Jones, and B. L. Webber, Eds.: *Readings in Natural Language Processing*. Morgan Kaufmann Publishers, Inc., Los Altos, CA (1986) 71-88.