Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Biomedical Informatics xxx (2006) xxx–xxx
Journal of Biomedical Informatics
www.elsevier.com/locate/yjbin

Informed consent procedures: An experimental test using a virtual character in a dialog systems training application

Robert C. Hubal [a,*], Ruth S. Day [b]

[a] Technology Assisted Learning Center, RTI International, Research Triangle Park, NC 27709, USA
[b] Department of Psychology, Duke University, Durham, NC 27708, USA

Received 30 May 2005

Abstract

Researchers are generally trained to administer informed consent by studying approved guidelines, but still can fail to satisfactorily answer questions from potential participants. An application using a virtual character allowed novice participants to practice administering informed consent. This character was designed to behave as a potential participant for a study and asked many of the questions research participants typically ask, such as queries about the study itself, the sponsor, timing, selection procedures, confidentiality, voluntariness, benefits and risks, and contact information. The user responded to the character's queries as if speaking with a true potential research participant. The application was effective even after only brief usage. In a laboratory experiment, novice participants who practiced with the virtual character were later more effective in conducting informed consent interviews with a human interviewee than those who were trained only with written materials. Thus, simulated learning-by-doing improved informed consent skills. Implications for related health dialog applications are discussed.

Keywords: Informed consent procedures; Intelligent agents; Discourse processes; Effectiveness evaluation; Virtual reality

1. Introduction

Effective health dialogs, as with any communications, begin with established rapport. In a health setting, particularly where communications involve possible changes in patient health practices and outcomes, rapport is partly established by the gaining of informed consent. The goal of the present study was to assess the effectiveness of an agent-based dialog system for training informed consent techniques.

This work was funded through the US National Institutes of Health (NIH) Human Subjects Research Enhancement program (HSREP). The purpose of HSREP was to "strengthen oversight" of human subjects research at institutions that received significant NIH research support. At the first author's institution (RTI International), three projects were undertaken: increased automation of the Institutional Review Board (IRB) information services; an interactive multimedia training program regarding the protection of human subjects focused on issues of direct relevance to the types of research conducted at RTI; and, the focus of this work, a virtual reality (VR) application for enabling researchers to practice administering informed consent to potential research participants. At the second author's institution (Duke University), the effectiveness of the application was examined in a laboratory experiment comparing the ability of novice participants to conduct informed consent interviews who were trained using the application versus those trained only with traditional written materials.

Research personnel are generally trained to administer informed consent by studying relevant guidelines and becoming certified by an IRB. Even though they use an approved informed consent form for a given study, they still do not always answer questions from potential participants in a satisfactory manner [1]. The VR training

* Corresponding author. Fax: +1 919 541 6965.
E-mail address: rhubal@rti.org (R.C. Hubal).

1532-0464/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2005.12.006

application allowed users to practice on typical elements [2] of informed consent, such as providing all relevant information necessary concerning participation, promoting participant comprehension of relevant information, and ensuring participant voluntariness to consent.

In the application, a virtual character (also referred to here as an intelligent agent) acted as a potential participant in a generic study and posed typical queries regarding the study sponsor, the study's content, how participant selection was made, confidentiality, duration, and contact information. These queries were to be addressed by the user. Only if the user addressed them satisfactorily would the agent agree to participate. When the user failed to address a query satisfactorily, the application provided a hint, ended the current interview trial, and initiated a new trial.

The application did not expressly train for greeting, conflict resolution, de-escalation, obtaining cooperation, rescheduling, or parting, though all of those are components of health dialogs. The application was also not complete with regard to all topics of informed consent; for instance, the potential participant's capacity and authorization to respond [3,4] were not addressed. The application was not 'scaffolded' [5] with instructional content nor tutoring, as it would need if it were standalone. Finally, as with all simulations developed at RTI [6], the application was never intended to replace existing forms of training, but instead to augment existing training. However, since existing training involves strictly familiarization (primarily with written materials) and no learning-by-doing except on the job, it was hypothesized that learning-by-doing in a simulated environment would improve informed consent skills.

2. Advances over prior work

The work built on related intelligent agent applications developed at RTI. These related applications include a trainer for field survey interviewers learning to avert nonresponse [7], a trainer for telephone survey interviewers practicing obtaining cooperation [8], a trainer for health care providers to interact with pediatric patients [9], and a tool for prevention researchers assessing conflict resolution skills demonstrated by at-risk adolescents [10]. Also, a preliminary study evaluated a prototype informed consent training application [11].

The related work involved some assessment of the design, development, usability, and acceptance of intelligent agent applications, but did not assess effectiveness. The work described here contributes both an approach for evaluating virtual character applications and an understanding of the usefulness of a particular application for informed consent.

3. Application components

3.1. Virtual environment and character

The virtual environment was a suburban/country kitchen scene, as would be typical for field interviews (i.e., household research) of any kind, including those involving health dialogs.

The agent was a virtual character, a 30-something, pleasant-looking woman of average build typical of actual household respondents. No special movements (animations) were needed beyond seated conversational gestures.

Fig. 1 illustrates the user interface, showing the virtual character and environment.

The application was rendered using an off-the-shelf gaming engine (NDL's Gamebryo) onto a standard monitor and ran on a personal computer. The PC had decent but not exceptional on-board and graphics memory (Pentium 4, Windows XP, 512 Mb main memory, NVidia GeForce Fx Go5200 video card).

Details about the development of the informed consent application can be found in [12]. Details about the virtual character architecture can be found in [11].

3.2. Natural language component

Among the needs of any health dialog application is a natural language component. For this work, the language component was implemented with an interpreter program following a transition network that was augmented with state and goal-oriented variables, context, and application-specific conditionals [13].

Specifically, grammar files were developed to feed into an off-the-shelf speech recognizer (IBM VoiceType), from which output was fed into a parsing algorithm. Rules in the grammar files broke down input sentences into expected phrases and words, then returned semantic or meaningful tags. For instance, the set of rules:

START → BRIEF_CONV : wont_take_long.
BRIEF_CONV → i ONLY have a few questions to ask you.
BRIEF_CONV → i ONLY want to speak with you for A_BRIEF_TIME.
BRIEF_CONV → this will ONLY take A_BRIEF_TIME.
ONLY → just.
ONLY → only.
A_BRIEF_TIME → a few minutes.
A_BRIEF_TIME → a short while.

would allow the parser to correctly interpret user responses such as "I only have a few questions to ask you.", "I only want to speak with you for a short while." and "This will just take a few minutes.". (The parser is actually able to recognize responses that do not exactly match any combination of rules, returning a confidence score along with a semantic tag; for details, see [13].) For any of those inputs, the parser would return a tag (wont_take_long) to a running interpreter. This interpreter would then, as described in the next sections, determine if any such sentence (i.e., implying that the interview would not take long) could be appropriate given the current state of the conversation.

Fig. 1. Virtual environment and character.

Natural language enabled the agent to adapt or tailor responses to the user's input. Language models that underlay the VR character's behavior were developed through many iterations of testing by subject-matter experts. This permitted building up grammars (there were some 1150 rules) to yield acceptable content for user responses.

3.3. User input: Tracked variables

Each query posed by the potential participant in a health dialog sets up expectations. For instance, if the character states "I am too busy.", then that sets up an expectation of any of several satisfactory responses by the user: "I only have a few questions to ask you.", "The study won't take a long time to conduct.", or "You represent a segment of the general population and your responses are very important to the integrity of the study." The interpreter transitioned to new conversational states modeled in its network based on the semantic tag of the user's response returned by the parser.

Tracked variables from the user's input, based on expectations, also guided the agent's behavior. Variables that were tracked included completeness (Was full information given in the user's response?), complexity (Did the response use long sentences, difficult words, or jargon?), responsiveness (Was the content of the agent's query addressed?), and truthfulness (Was the content addressed accurately?). For example, if the character asked "Do I have to participate?", the answer "Yes." is responsive but false, the answer "No." is responsive and truthful but vague or incomplete, the answer "Your participation is completely voluntary." is responsive, truthful, complete, and reassuring, and the answer "I only have a few questions to ask you." is not responsive, vague, and not particularly reassuring.

3.4. Conversational flow diagrams

Conversation flow diagrams were developed with subject-matter expert input for the following components of informed consent: benefits and compensation, confidentiality, contacts, study duration, the research, selection methods, and voluntariness. These diagrams guided the agent's verbal behavior; the diagrams specified what valid and invalid user responses to expect to any given query and how to respond in turn to that user input. Fig. 2 shows the flow diagram for the script following the character asking "Do I have to participate?". (Another flow diagram can be found in [12].)

3.5. Hint table

A hint table was constructed that matched each VR character query with expected user responses. The informed consent interview terminated on bad or incorrect responses, whether predetermined (i.e., expected) or not;

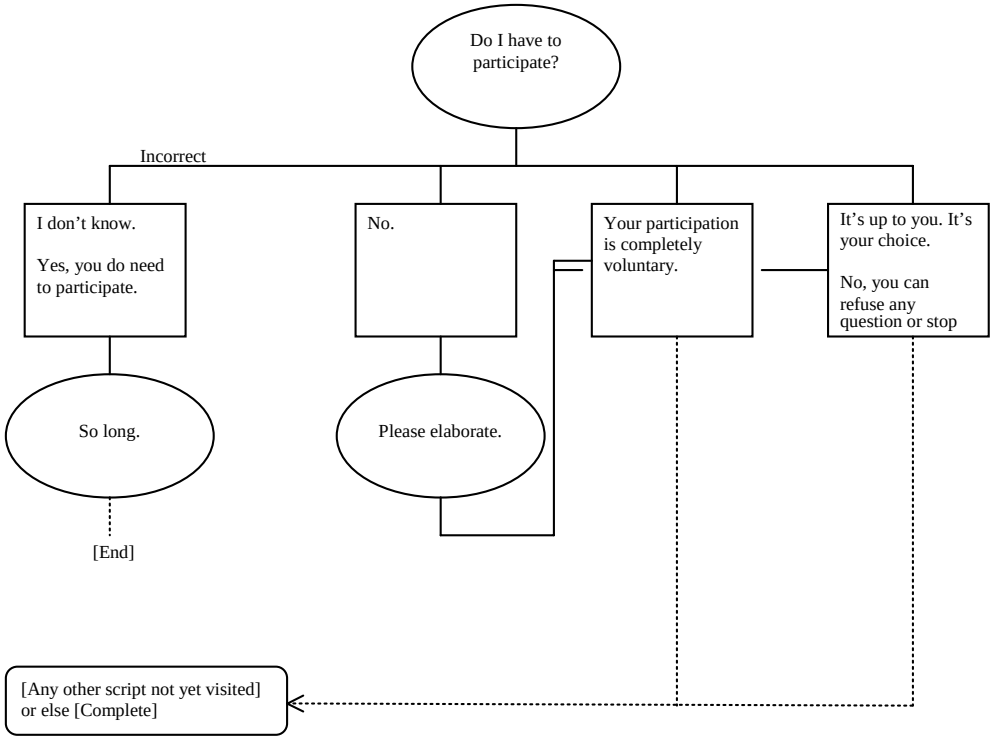

Fig. 2. Voluntariness script.

for predetermined incorrect responses, the application displayed an appropriate hint previously validated by experts. To succeed with the entire conversation (about nine conversational turns), the user was required to complete one full interaction, giving satisfactory responses to all questions. Returning to the "I am too busy." example, if the user instead responded with a description of the study, a hint was brought up via a pop-up box noting that the character was at that moment concerned about how long the research would take, not about the research effort itself. Or, if the user instead insisted that the character had to participate in the study, then a hint was brought up stating that no potential participant has to participate in a research study (see Fig. 3). (A portion of the hint table can be found in [12].)

4. Methods

4.1. Participants

The participants were 24 undergraduate students at Duke University, fulfilling a research requirement for an introductory psychology course. They had no particular experience with health dialogs in general or informed consent in particular, beyond what might be expected of undergraduate students. All were fluent in English. On a random basis, half were assigned to the VR (experimental) condition and half to the control condition (described below). Their average age was 19.0 years, with a nearly even split between males and females (13 vs. 11, respectively). Participants were told this was a study on learning and developing interview skills.

4.2. Written training materials

A three page booklet served as written training materials, entitled "Health Registry Survey—Informed Consent Interviews". It described a research study on the effects of environmental pollution on health. Most of the information was on how to obtain informed consent from potential research participants for this hypothetical study. Topics included how participants were selected, how much time the survey would take, types of questions to be asked, benefits and risks, contact information, privacy of information collected, and other typical informed consent information. The booklet was just over 1200 words long and was written at the 6.7 grade level (as determined by the Flesch–Kincaid index). (The Flesch–Kincaid grade level index estimates how many years of schooling a reader would require to understand the content. It includes both word familiarity and sentence length in its calculation.)

4.3. Procedure

Each participant was tested individually in a cognition laboratory, in a sound-shielded testing booth. The experiment session consisted of four main phases:

4.3.1. Study phase

Participants studied the written materials about informed consent. They were told that the booklet con-

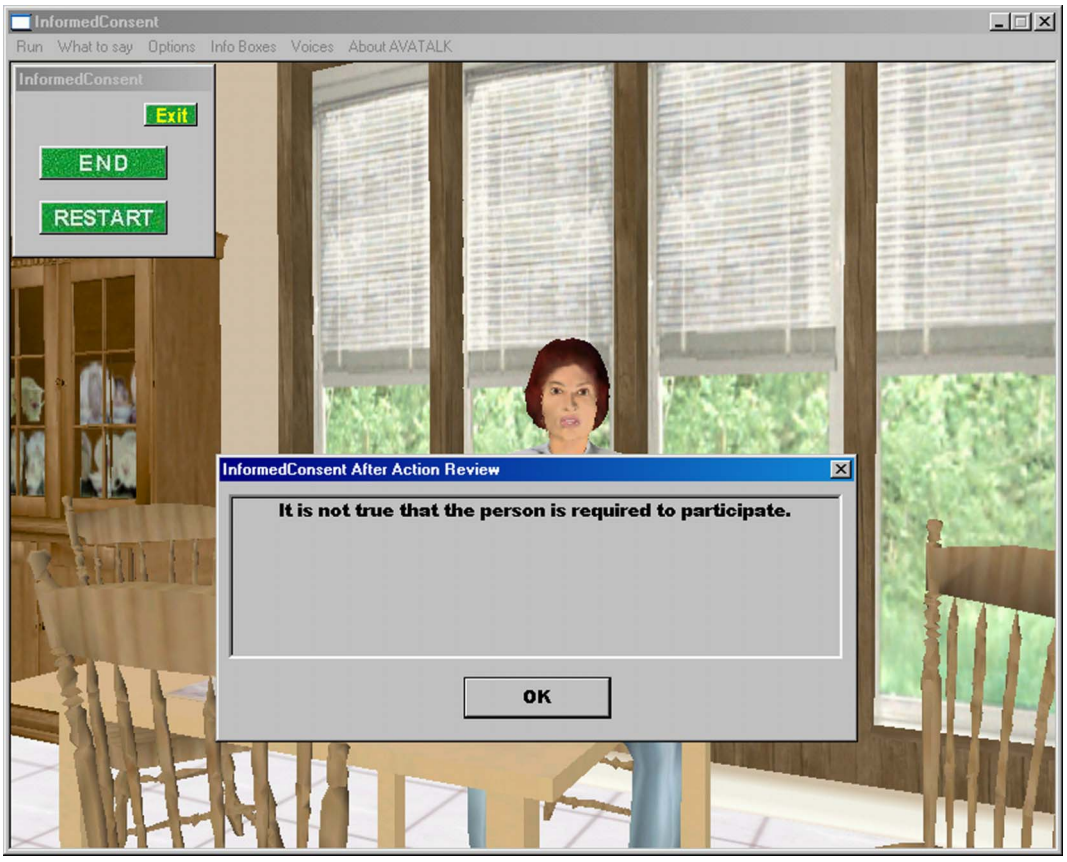

Fig. 3. Hint displayed on bad user response.

tained information about interviewing potential participants for a health survey. Some of this information they might tell the potential participant, while some of it might be relevant to questions the person might ask. Therefore, it was to be "used as a flexible guide, rather than [as] a script to be memorized".

Participants were asked to study the information for the full amount of time given (8 min), so that they would be able to conduct such an interview. Those in the VR condition then went on to the VR training phase (see below), while those in the control condition kept studying for 12 more minutes. All participants were told that they would practice their interview skills later, in a question/answer format. Therefore, they were to think about the kinds of questions potential participants might ask while studying the booklet.

4.3.2. VR training phase

Participants in the VR condition practiced interviewing skills with the VR character. The character sat at a virtual table facing the human participant and asked questions such as, "Do I have to participate?" and "Will my answers be kept private?". The questions asked were a random subset from a pool of ten questions (with variant recordings for each question to reduce repetitiveness while retaining the meaning of the question). The participant answered the questions by speaking naturally into a microphone. Thus this was a non-immersive setting (i.e., the character appeared on a PC screen, with no special equipment or instrumentation required except for a microphone). The participant was encouraged to speak naturally, as if to a real person. Responses were captured by the software and were also audiotaped for later analysis.

Participants were given two major goals for the interviews, to answer the VR character's questions correctly based on the information they had just studied and to eventually gain her consent to participate in the health survey. If the participant answered a given question incorrectly or incompletely, the VR character said "Please repeat that." or "I am not convinced.". Sometimes she asked the same question again. Each interview concluded in one of two possible ways. If the participant answered all questions correctly, the VR character said "Let us go.", thereby agreeing to participate. If not, the character said "Goodbye." or "I am not interested.", a hint might be displayed, and that interview ended.

After concluding an interview, another interview was initiated, using a new subset and random order of the test questions. This training process continued for 5 min, during which the participant completed several interview trials. Pilot work revealed that VR training raised questions about the written material participants had just studied. Therefore, VR participants took a break from VR training and re-read the written materials for another 2 min, then

conducted additional VR interviews for five more minutes. The entire VR training time was 10 min. Given 2 min for instructions, the 12 min devoted to this condition matched the additional study time for the control participants.

4.3.3. Human interview phase

All participants—both those in the control and VR conditions—conducted an informed consent interview with a human interviewee. This mock interviewee was actually an experimenter who asked the same set of ten questions for all participants (see Table 1) according to two different random orders. Hence, the participants were required to apply the knowledge they had just gained by studying the written materials and (for those participants in the VR condition) practicing with the VR character. This part of the session was videotaped for later behavior coding.

4.3.4. Evaluation phase

At the end of the experiment session, participants in the VR condition evaluated the usefulness of the written material, human interview, VR interviews, and overall training session. Then they rated how much they liked working with the written information, human interviewee, and VR character. They also rated features of the VR character, including how realistic she was as a person, her speech, and body movements. The list of evaluation questions is shown in Table 2. Scales for these ratings are shown in Table 3. Control participants performed the same ratings, excluding those regarding the VR system.

4.4. Behavior coding

For the final human interviews, two independent behavior coders, knowledgeable about informed consent issues and aware of the experiment but blind to participants' test condition (experimental or control), rated participant responses according to the scale in Table 4. Eight ratings measures were collected for each participant by each coder such as the acceptability of participants' responses, the appropriateness of participants' reaction time, and overall realism of the conversation, as shown in Table 5. The time taken by each participant to conduct the entire interview was also measured.

Table 1
Questions asked in the human interview

How long will the survey take?
Will my responses be kept confidential?
What will be on the survey?
Do I have to participate?
Will I get anything from the study?
What is the registry?
Who is conducting this study?
Who do you work for?
Who can I call if I have questions?
Do I have to answer all questions?

Table 2
Evaluation and liking questions asked at the end of the experiment

Overall, how would you rate the usefulness of:
The written booklet
The human interview
The computer interviews[a]
The entire interview training session
Computer character:[a]
Overall, how realistic was the computer person?
How would you rate her speech?
How would you rate her body movements?
How well did you like working with:
The written booklet
Interacting with the computer character[a]
The final interview with the human
Do you have any comments about the session?

[a] Participants in the control condition were not asked about the virtual character.

Table 3
Participant evaluation scales

Evaluation scale
5 = excellent
4 = good
3 = moderate
2 = fair
1 = poor
Liking scale
5 = like very much
4 = like
3 = neutral
2 = dislike
1 = dislike very much

Table 4
Behavioral coding scale

5 = extremely
4 = very
3 = somewhat
2 = not too
1 = not at all

4.5. Statistical analyses

Ratings for participant behavior, time taken to answer all human interviewee questions, and perceived training value were analyzed using a general linear model for effects of random order of questions and test condition (experimental vs. control). Pearson correlation as well as κ coefficients were calculated to compare behavior coding between the independent raters. A further analysis of variance examined the possible effect of the experimental (VR) condition on behavior coding and participant ratings, as well as time taken by participants during the final human interview.

Table 5
Behavioral coding measures for the final human interview

(1) In general, how acceptable would you rate the person's responses to questions or concerns posed by the potential respondent?
(2) In general, how appropriate would you rate the response time of the person in answering questions or addressing concerns?
(3) How well do you believe the person comprehended what question was being asked or concern was being raised?
(4) How well do you believe the person "thought on his/her feet"?
(5) How well did the person moderate his/her tone of voice?
(6) How appropriate do you feel this person's body language was?
(7) How successful do you think this person would be at gaining respondent cooperation?
(8) In general, how realistic did you find the overall conversation between the persons?

5. Results

5.1. Initial analyses

The random order of test questions in the final human interview did not influence the behavior coders ($F < 1$) and was omitted from further analyses.

The correlation between independent behavior coders was quite high and was highly significant for all of the eight measures shown in Table 5, ranging from $r = 0.58$ to 0.81 (all $p < .004$). A highly significant unweighted κ of 0.31 also demonstrated agreement between coders. Coder responses were combined and averaged across all eight measures for subsequent analyses.

5.2. Main results

Participants in the experimental condition conducted an average of 6.7 (SD = 1.4) interviews with the VR character during the ten minute training period. These interviews averaged 7.1 conversational turns each (SD = 1.6, range = 1–18), suggesting rapid learning since each interview trial ended upon a non-satisfactory response to the VR character's question.

Experimental participants' rating of the human interview was significantly higher than that for control participants (4.3 vs. 3.3, $p < 0.001$), suggesting that even though experimental participants were not particularly impressed with the overall realism of the virtual character (average = 2.8, SD = 1.0 on the evaluation scale shown in Table 3), and only moderately liked interacting with the character (average = 3.5, SD = 1.0), the practice with the character better prepared them to like the interview and find it useful.

Further analyses support this idea of the experimental participants being better prepared for the final interview. For ease of presentation, three summary variables were used for comparing the experimental versus control conditions: (1) participant behavior in the final interview, averaged across both coders and all questions, (2) time taken by the participant to respond to all of the human interviewee's questions during the final interview, and (3) participant rating of the overall training value of the session. Table 6 shows values for these variables. These variables were chosen because they pertain to all participants and because they reflect three types of measures, namely independent coding of participant behavior, an objective measure (response time) of participant behavior, and a subjective measure provided by the participant.

Table 6
Summary statistics for variables of interest

Variable	Mean	N	Condition
Participant behavior: combined behavior coder ratings averaged across all eight items	3.1	11[a]	Control
	4.0	12	Experimental
	$F(1,21) = 5.61$, $p < 0.03$		
Time taken: average time (s) taken by participants to respond to all questions posed during final human interview	107	11[a]	Control
	89	12	Experimental
	$F(1,21) = 1.95$, n.s.		
Training value: average participant rating of session training value	3.6	12	Control
	4.0	12	Experimental
	$F(1,22) = 2.57$, n.s.		

n.s., not significant.
[a] The video data for one control participant was unavailable to the independent behavior coders.

All summary variable values were in the expected direction, and the behavior coding ratings were significantly different between conditions. That is, participants in the experimental condition tended to require less time to respond to questions posed during the final human interview and to rate the learning session as more useful than participants in the control condition. Independent observation by the two behavior coders revealed that experimental participants were significantly better than those in the control condition at answering questions posed by the human interviewee during the final interview. (All eight ratings in fact differed significantly between conditions, so that experimental participants were rated as better comprehending the interviewee's question, responding more appropriately to the question, and more likely to obtain cooperation from the interviewee, than control participants; see Table 7.) Since the only methodological difference between participants in the experimental and control conditions was VR practice time at the expense of study time, the findings suggest that simulated learning-by-doing improved informed consent skills.

An analysis of experimental participants' dialog with the VR character clarifies why the practice would have helped. The participants were able to successfully complete one-quarter of trials (20 complete out of 80 total) during practice with the VR character (i.e., reach a point after sufficient conversational turns where all of the character's informed consent questions are answered and where the

Table 7
Average behavioral coding measures for experimental and control participants

Measure	Control group	Experimental group	Significance (t test)
Acceptable response	3.0	3.8	$p < 0.04$
Reaction time	3.5	4.3	$p < 0.02$
Comprehension	3.1	4.2	$p < 0.02$
Thinking on the feet	2.7	3.8	$p < 0.01$
Moderating the voice	3.6	4.3	$p < 0.05$
Using body language	3.0	3.8	$p < 0.02$
Likelihood of success	2.7	3.7	$p < 0.03$
Overall realism	3.2	3.9	$p < 0.04$

character agrees to participate). On average it took them up to their fourth trial before achieving success. For the remaining trials the character interpreted the participant's response as not accurate, and ended the trial without success, on occasion (when it was a subject-matter expert predetermined incorrect response) providing a hint.

These trial-ending conversational turns, though, represented only 15% of all conversational turns between the participant and VR character. Another 7% of the time the character failed to understand the participant and asked for the participant to repeat his/her response. The remainder of the turns—over three-quarters of them, some 416 conversational turns—involved either the participant practicing by providing a direct response to an informed consent question posed by the character, or the character prompting the participant to clarify or expand on his/her response or provide another satisfactory response.

6. Conclusions

These results have implications for many types of health dialog systems. Informed consent is essential for health services research [14] and precedes all other healthcare provider/patient communication. The informed consent process helps establish trust between the researcher and participant or between the healthcare provider and patient. This work showed that a VR training application can improve dialog performance for participants with no previous interview training. Further, informed consent is critical in many fields, but particularly in medicine, where informed consent is a component of professionalism, one of six general competencies for residency programs advocated by the American Accreditation Council for Graduate Medical Education [15], and a vital component of clinical trials research.

The methods used in this work to develop and test the application follow procedures used elsewhere by the authors and others [16–19]. Application methods included: design of the virtual environment with, in this case, one appropriate character in an appropriate setting; subject-matter expert validated scripts guiding the agent's verbal behavior and response expectations; tracking how well the user's responses followed the agent's queries; compilation of a hint table to pair expected incorrect responses with hints; and extensive expert user testing that involved logging their dialogs and revising language grammars to incorporate their input and ensure the agent responded appropriately. Additional methods for testing application effects on users included: developing companion written materials, designing a laboratory test session, evaluating participants' interview performance, and obtaining participants' views about the training experience. The results of evaluation of participants' interview performance supplement findings from [10] that showed that observer ratings of these same measures of participant behavior demonstrate construct and criterion validity with established measures of psychosocial factors and behavioral criteria.

A limitation to the conclusions from the present study centers on the use of only one group of experimental participants. That is, the benefit seemingly produced by practice with a virtual character may derive simply from the act of practicing rather than studying (as was requested of participants in the control condition). It is possible that an embodied virtual character is not needed, but that practice with a spoken dialog system is sufficient to achieve the same benefits.

Partial evidence to support this assertion comes from a study that evaluated the use of a disembodied virtual character for simulating the environment a telephone interviewer faces during the opening of an interview [8]. In that study, some three-quarters of trainees perceived an increase in their ability to deal with questions or concerns raised by the interviewee, and nearly all felt the application helped them to some degree to think on their feet, although their actual performance was not assessed.

The added training benefit of embodying virtual characters within a spoken dialog system, then, remains in question [20], though it is likely that users' experiences (as measured by enjoyment or presence) would increase [21–23]. A follow-up experiment is planned to control for embodiment and attempt to isolate training benefits.

As mentioned, the present application was not designed for training greeting/parting, conflict resolution, de-escalation, obtaining cooperation, or rescheduling. Given the success of this application and others (e.g., [8,10,16]), future applications addressing these components of health dialogs are warranted. For all of the components, the content of dialogs can be reasonably well defined so that the natural language and tracking of user responses can be formalized.

This is one of few studies to test the effectiveness of training with VR characters, using experimental methods developed in a cognition laboratory. Most previous studies relied on user acceptance or usability of applications [24,25], rather than their effectiveness. Intelligent agent simulations, it has been argued (e.g., [6]), can improve interaction skills training and assessment by providing students with more practice time and consistent interaction experiences. This is true of simulations in general, where students can acquire and practice skills in a safe, reliable, modifiable environment. This work lends support for the use of intel-

ligent agent simulations to train and assess informed consent skills in particular, and possibly health dialog systems more generally.

Acknowledgments

The authors thank the following individuals for their assistance in developing the intelligent agent application and conducting this research: Curry Guinn, Henry Schwetzke, Wendy Visscher, and Kristen Zukosky. This research was supported by Grant No. EIA-0121211 from the US National Science Foundation (NSF) and Grant No. 2 S07 RR018257-02 from the US National Institutes of Health (NIH). Opinions expressed in this document are those of the authors, and do not necessarily represent the official positions of NSF, NIH, RTI International, or Duke University.

References

[1] Sugarman J, Lavori PW, Boeger M, Cain C, Edson R, Morrison V, Yeh SS. Evaluating the quality of informed consent. Clin Trials 2005;2:1–8.
[2] Pedroni JA, Pimple KD. A brief introduction to informed consent in research with human subjects. Bloomington, IN: Poynter Center; 2001.
[3] Grisso T, Appelbaum PS. Comparison of standards for assessing patients' capacities to make treatment decisions. Am J Psychiat 1995;152(7):1033–7.
[4] Bean G, Nishisato S, Rector NA, et al. The assessment of competence to make a treatment decision: an empirical approach. Can J Psychiat 1996;41:85–92.
[5] Jackson SL, Krajcik J, Soloway S. The design of guided learner-adaptable scaffolding in interactive learning environments. CHI 1998:187–94.
[6] Hubal R, Guinn C. Interactive soft skills training using responsive virtual human technology. In: Proceedings of the interactive technologies conference. Arlington, VA: Society for Applied Learning Technologies; 2003.
[7] Camburn DP, Gunther-Mohr C, Lessler, JT. Developing new models of interviewer training. In: Proceedings of the international conference on survey nonresponse. Portland, OR; 1999.
[8] Link MW, Armsby PP, Hubal RC, Guinn CI. Accessibility and acceptance of responsive virtual human technology as a survey interviewer training tool. Comput Human Behav 2006;22(3):412–26.
[9] Hubal RC, Deterding RR, Frank GA, Schwetzke HF, Kizakevich PN. Lessons learned in modeling virtual pediatric patients. Stud Health Technol Inform 2003;94:127–30.
[10] Paschall MJ, Fishbein DH, Hubal RC, Eldreth D. Psychometric properties of virtual reality vignette performance measures: a novel approach for assessing adolescents' social competency skills. Health Educ Res 2005;20(1):61–70.
[11] Guinn C, Hubal R, Frank G, Schwetzke H, Zimmer J, Backus S, et al. Usability and acceptability studies of conversational virtual human technology. In: Proceedings of the SIGdial workshop on discourse and dialogue. East Stroudsburg, PA: Association for Computational Linguistics; 2004.
[12] Hubal RC, Guinn CI, Sparrow DC, Studer EJ, Day RS, Visscher WA. A synthetic character application for informed consent. In: Proceedings of the workshop on dialogue systems for health communication. New York: ACM Press; 2004.
[13] Guinn C, Hubal R. Extracting emotional information from the text of spoken dialog. In: Proceedings of the workshop on assessing and adapting to user attitudes and affect: why, when and how? Johnstown, PA: User Modeling, Inc.; 2003.
[14] Health Services Research and the HIPAA Privacy Rule. NIH Publication Number 05-5308. Washington, DC: US Department of Health and Human Services 2005.
[15] <http://www.acgme.org/outcome//>.
[16] Frank G, Guinn C, Hubal R, Pope P, Stanford M, Lamm-Weisel D. JUST-TALK: an application of responsive virtual human technology. In: Proceedings of the interservice/industry training, simulation and education conference. Arlington, VA: National Training Systems Association; 2002.
[17] Bickmore T, Cassell J. Social dialogue with embodied conversational agents. In: van Kuppevelt J, Dybkjaer L, Bernsen N, editors. Natural, intelligent and effective interaction with multimodal dialogue systems. New York: Kluwer Academic; 2004.
[18] Johnson WL. Interaction tactics for socially intelligent pedagogical agents. IUI 2003:251–3.
[19] Marsella S, Gratch J. Modeling coping behavior in virtual humans: don't worry, be happy. AAMAS 2003:313–20.
[20] Nijholt A. Agents, believability and embodiment in advanced learning environments. ICALT 2001:457–9.
[21] Lester JC, Converse SA, Kahler SE, Barlow ST, Stone BA, Bhogal R. The persona effect: affective impact of animated pedagogical agents. CHI 1997:359–66.
[22] Maldonado H, Hayes-Roth B. Toward cross-cultural believability in character design. In: Trappl R, Payr S, editors. Agent Cultures. Mahwah, NJ: Lawrence Erlbaum and Associates; 2003.
[23] Prendinger H, Ishizuka M. Life-like characters: tools, affective functions and applications. Berlin: Springer-Verlag; 2004.
[24] Marriot A, Pelachaud C, Rist T, Ruttkay S, Vilhjálmsson H, editors. Proceedings of the workshop on embodied conversational agents—let's specify and evaluate them! AAMAS 2002.
[25] Ruttkay Z, Dormann C, Noot H. Embodied conversational agents on a common ground. In: Ruttkay Z, Pelachaud C, editors. From brows till trust: evaluating embodied conversational agents. Dordrecht: Kluwer; 2004.