Embodied Conversational Agents to Promote Health Literacy for Older Adults

A. SPECIFIC AIMS

We propose to develop and evaluate technologies to compensate for inadequate or marginal functional health literacy in older adults. An embodied conversational agent—a computer character that simulates face-to-face conversation using voice, hand gesture, gaze cues and other nonverbal behavior—will be used in a clinic-based kiosk to review written materials given to patients with Type 2 diabetes mellitus. To this end, we will develop the conversational agent kiosk as a laboratory tool to be used in a series of experiments to evaluate the efficacy of different media and verbal and nonverbal conversational behavior for knowledge acquisition by older adults with low literacy skills.

The **specific aims** of this project are to: (1) construct the baseline conversational agent kiosk system; and (2) evaluate the hypotheses described below in a pilot study. The study will compare diabetes education provided by a conversational agent to the standard of care for diabetes education in most medical practices (diabetes education pamphlet distribution), and evaluate the efficacy of each on immediate (immediately following the intervention) and distal (at three-month follow-up) diabetes knowledge acquisition by participants. Increases in knowledge of diabetes is also expected to lead to better glycemic control (decreases in HbA_{1c} values between the time of intervention and followup). Functional health literacy level is expected to play a moderating role on the efficacy of the interventions.

The following **hypotheses** will be evaluated:

1. Immediate and distal knowledge gains and glycemic control will be significantly improved when the current standard of care is augmented with a brief "virtual consultation" with an embodied conversational agent, compared to the current standard of care alone, and the improvement will be most pronounced in patients with inadequate or marginal functional health literacy.
2. Patient satisfaction will be greater when the current standard of care is augmented with an embodied conversational agent, compared to the current standard of care alone.

This research will advance our understanding of communication processes in older adults with low literacy skills, and provide specific techniques for ensuring comprehension of complex health information. This work will also advance our understanding of how older adults can effectively use computer technology to obtain health information. Contributions to the fields of human-computer interaction and artificial intelligence include knowledge of how older adults accept and interact with conversational agents in health care settings, how their use of nonverbal conversational behavior differs from the behavior of younger adults, and how student models in intelligent tutoring systems need to be adapted for health communication with older adults with low literacy skills.

B. BACKGROUND AND SIGNIFICANCE

B.1 Type 2 Diabetes Mellitus in Older Adults

Diabetes is a major source of morbidity, mortality, and economic expense in the U.S. The overall prevalence of diabetes in the U.S. is 4.0%, climbing to 13.1% and 13.9% for white men and women aged 65-74, respectively, and 19.2% and 29.9% for black men and women aged 65-74.[1] The overall prevalence is expected to increase by 165% over the next 50 years, but for black men and women over 75, the projected increases are 352% and 555%, respectively.[1] People with diabetes are at risk for the development of specific acute metabolic complications, such as diabetic ketoacidosis, hyperglycemic hyperosmolar nonketotic coma, and hypoglycemia, and a variety of chronic complications specific to diabetes, including circulatory, renal, ophthalmic, neurological, and skin disorders.[2] Direct medical and indirect expenditures attributable to diabetes in the U.S. in 1997 were estimated at $98 billion.[2]

B.2 Functional Health Literacy in Older Adults and in Patients with Diabetes Mellitus

More than one-third of U.S. adults over 65 have inadequate or marginal functional health literacy, and among indigent and minority patients in urban areas this number rises to over 80%.[3] Functional health literacy—the ability to perform the basic reading and numerical tasks required to function in the health care environment—affects patients' ability to understand medication labels and instructions, hospital discharge instructions, instructions for assistive devices and medical equipment, and educational material.[4] Patients with inadequate health literacy report lower health status, [5, 6] are less likely to use screening procedures, follow

medical regimens, keep appointments, or seek help early in the course of a disease,[7] have greater difficulties naming their medications and describing their indications,[8] more frequently hold health beliefs that interfere with adherence,[9] have higher health-care costs,[7] and have higher rates of hospitalization.[5] Improving health literacy is one of the stated goals of Healthy People 2010.[10]

Diabetes patients with inadequate or marginal functional health literacy demonstrate significantly less knowledge about their disease than diabetes patients with adequate literacy.[11] For example, a recent study of patients with diabetes found that 94% of those with adequate health literacy knew the symptoms of hypoglycemia, compared with only 50% of those with inadequate literacy.[11] Inadequate health literacy is also associated with worse glycemic control and higher rates of retinopathy in patients with Type 2 diabetes mellitus.[12] Given the exceedingly high prevalence of Type 2 diabetes in older adults, the high prevalence of inadequate functional health literacy in older adults, including diabetes patients, and the association of reduced health literacy in diabetes with adverse effects on diabetes outcomes, maximizing understanding of and adherence to the complex treatment regimens for diabetes represents an important priority in the care of these patients.

B.2.1 Addressing Inadequate Health Literacy

Although low literacy in early adulthood is a significant contributor to inadequate functional health literacy in older adults, a number of additional factors contribute to the problem in older adults. Older adults show age-related declines in a variety of linguistic domains—including syntax and semantics, and lexical retrieval and discourse—as well as in certain cognitive abilities that can impact linguistic ability, including inhibitory efficiency, reduced processing speed, and deficits in working memory. [13] Changes in sensory abilities such as visual perception and hearing as well as changes in affect, motivation, and social status also may play a role in declining linguistic ability.[13]

…

B.3 Embodied Conversational Agents

Embodied Conversational Agents (ECAs) are animated computer-based characters that use speech, gaze, hand gesture, intonation and other nonverbal modalities to emulate the experience of human face-to-face conversation with their users.[14] Such agents can provide a “virtual consultation” with a simulated health provider, offering a natural and accessible source of information for patients in general, but especially older patients with low literacy skills, and a low-pressure environment in which patients are free to ask questions and take as much time as they need to understand the information they require.

In addition to the systems described in Section C, some of the other major ECAs developed to date are Steve,[15] the DFKI Persona,[16] Olga,[17] Gandalf,[18] SAM,[19] MACK,[20] and various pedagogical agents.[21-24] There are also a growing number of commercial ECAs, such as those developed by Extempo, Headpedal, and Artificial Life, and the Ananova newscaster developed by Ananova, Ltd. These systems vary greatly in their linguistic capabilities, input modalities (most are mouse/text/speech input only), and task domains, but all share the common feature that they attempt to engage the user in natural, full-bodied (in some sense) conversation. Microsoft has also produced a toolkit for developing animated talking agents (Microsoft Agent), although these characters are unable to use speech and nonverbal modalities at the same time, making them unusable for natural, multi-modal conversation.

…

B.3.1 Reasons for Using Embodied Conversational Agents for Health Literacy

There are several reasons why embodied conversational agents could provide an effective medium for patient education targeted at older adults with low health literacy. First, the human-computer interface relies only minimally on text comprehension (or not at all, in the case of speech recognition) and uses the universally understood format of face-to-face conversation, thus making it less intimidating and more accessible for patients with low literacy skills. In addition, as cited above, one study of pedagogical agents compared information delivery to students via an ECA that used speech output with an identical system that used text output instead, and found that students recalled more in the speech condition.[23]

…

B.4 Significance

Functional health literacy is a relatively new field of inquiry, but one with potentially significant impacts on public health, especially for older adults, minorities, and patients with low levels of education or socioeconomic status. Few proven methods exist for providing effective health education for these populations, thus policy in this area is limited to directives to provide patients with simplified written materials. The studies described in this proposal will provide important information about how comprehension of health materials can be maximized in these populations, as well as additional data on the prevalence of functional health literacy problems and their impact on health knowledge and disease outcomes. If successful, data from the studies should also highlight areas of research for future funding to help eliminate health disparities: the overarching goal of Healthy People 2010.

B.4.1 Clinical Contributions

The proposed development effort will yield a health education conversational agent that will be a useful tool for geriatric clinicians to use to educate their diabetes patients about proper self-care techniques, especially those with low literacy skills. The kiosk-based system could also readily be applied to other areas of health education for older adults, and the techniques developed and refined will also be applicable to home-based and Internet-based educational and behavior change systems. The series of proposed studies should also yield specific health communication techniques that could be used in human health provider-based patient education programs.

B.4.2 Scientific Contributions

The proposed research program will contribute to knowledge regarding automated methods for teaching complex information to older adults with low literacy skills. Information will be obtained on the usability of touch screen input vs. speech recognition for conversational interfaces. The acceptance and perception of ECAs by older adult patients will also be investigated.

Finally, this research will contribute to our basic understanding of communication processes in older adults and how these vary with functional health literacy skills. Specific areas of investigation include the use of verbal and nonverbal behavior by older adults to indicate understanding or misunderstanding of verbally-delivered content, and the prevalence, form and function of hand gestures and gaze behavior in face-to-face conversation with a health provider. This work will also make important contributions to the fields of human-computer interaction and embodied conversational agents, by providing information on the efficacy of different interface modalities and verbal and nonverbal communication strategies.

C. PRELIMINARY STUDIES

… Description of previous related work by the investigators …

D. RESEARCH DESIGN AND METHODS

The objective of the proposed research is to extend our previous work in the domain of embodied conversational agents into the domain of health information communication to older adults with low literacy levels. The research will also investigate how elders with low literacy levels communicate in health care settings. To this end, a kiosk-based embodied conversational agent will be developed that will be designed to help compensate for inadequate functional health literacy by reviewing written materials with older adult patients diagnosed with Type 2 diabetes. The kiosk will serve as a tool that will be used in a series of three randomized pilot studies to determine its efficacy, and the specific efficacy of different verbal and nonverbal conversational strategies that can be incorporated into the agent. The kiosk research tool will also be used to study how older persons with inadequate functional health literacy communicate in a controlled setting in which important health information is being communicated to them.

Each of the three sequential studies described in sections D.2-D.6 follow the same general plan and design. Experiment 1—described in section D.2—is described in complete detail, thereafter only differences from this design are described for each of the subsequent studies.

D.1 Project Timeline

The timeline for the project is presented in Figure 5. Development of the baseline intervention model and pilot testing will be developed during months 1-23. The experiment, conducted to assess the efficacy of the baseline model compared with a non-intervention control, will be conducted in months 25-36.

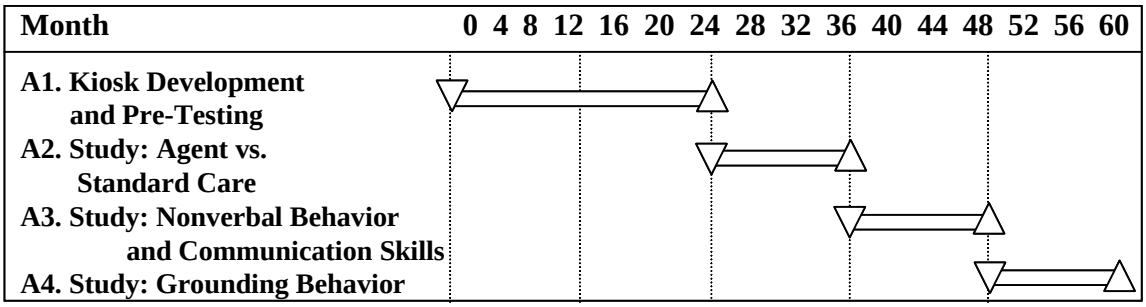

Figure 5. Project Timeline

D.2. Specific Aim 1: To Develop a Kiosk-based Embodied Conversational Agent

D.2.1 Technology Development

The embodied conversational agent developed for the MIT FitTrack system (described in Section C.4) will be adapted for the kiosk. The FitTrack agent uses a standard vector-based drawing engine (Macromedia Flash) coupled with a standard text-to-speech voice synthesizer (Microsoft Speech API) to provide a general-purpose conversational agent that will run on a very wide range of PC Windows-based platforms.

…

D.2.2 Current "Standard of Care" Education Pamphlet

As part of routine diabetes care in the Bohemia Medical Center (BMC) Geriatrics Clinic, all subjects (experimental and control) in all studies will be given a diabetes education pamphlet ("Managing Your Diabetes", published by Eli Lilly and Company). …

D.2.3 *Pre*-Testing

Extensive pre-testing of the system will be conducted in the first two years of the effort, using 30 subjects in Year 01 and 60 subjects in Year 02. The protocol for identifying these pretest subjects is identical to the protocol used to obtain the primary subjects, described in Section D.3.4.2.

A set of pre-test studies will be conducted with clinicians from the Geriatric Aardvark Practice, as well as Drs. Dewey, Cheatem, and Howe, to garner feedback on the correctness and appropriateness of the content delivered by the agent, as well as feedback on all aspects of the experimental set up and design.

A second set of pre-test studies will also be conducted as usability studies, using older adults representative of our target population, to ensure that agent's speech and all displayed text is understandable, and the touch screen and speech recognition result in acceptable input error rates. For some of the pre-testing, research staff will directly observe the subjects while they interact with the system. Subjects will be debriefed about their experience at the end of the interaction. Analysis of direct observation, interview and electronic data will be used to assess a number of issues about the system design and to make appropriate changes to the system, how it is used or how users are instructed to use it. Modifications to the system will be made as necessary to make the system as user friendly as possible.

D.3 Experiment 1: Randomized Pilot Study to Test the Efficacy of ECA vs. a Non-Intervention Control

The purpose of Experiment 1 is to determine if the immediate and distal knowledge gains and glycemic control will be significantly improved by a brief "virtual consultation" with an embodied conversational agent (BASELINE-AGENT), compared to the current standard of care alone (CONTROL), and to determine if the improvement is most pronounced in patients with inadequate or marginal functional health literacy.

In the following sections we describe the research model, the sample and sampling procedures, and the analysis plan.

D.3.1 The Research Model

The research model, Figure 6, defines the variables to be measured and their proposed associations. Participants are assessed at three time points during each experiment: T0, a baseline pre-intervention assessment; T1, immediately following the intervention; and T2 at a three-month follow-up.

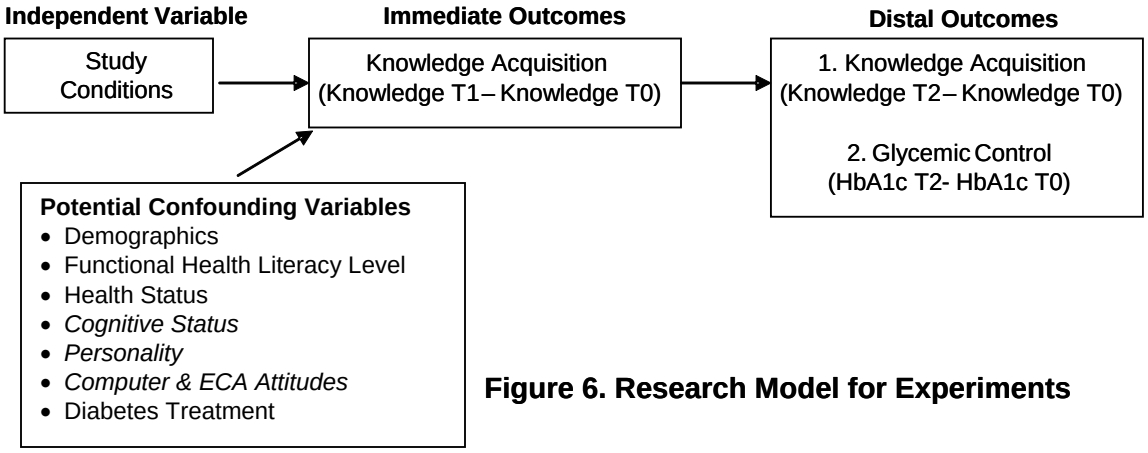

Figure 6. Research Model for Experiments

D.3.1.1 Independent Variables

The independent variables are the kiosk-based Embodied Conversational Agent (Experimental Condition) vs. a non-intervention control condition. The experimental and control conditions are described in Section D.3.

D.3.1.2 Primary Dependent Variables

Diabetes Knowledge. Diabetes knowledge will be assessed using the Diabetes Knowledge (DKN) Scales, three separate 15-item multiple choice questionnaires that measure general diabetes knowledge.[25] Reliability for the items in the scales (Cronbach's alpha) was 0.92, indicating high internal consistency. Validity was assessed by determining that 219 participants who participated in a 1-1/2 day class on diabetes scored significantly higher posttest on the measures compared to pretest (11.27 vs. 7.61, $p<.001$).

We will administer the DKN immediately before the educational intervention (T0), immediately following the intervention (T1), and at three months follow up (T2).

HbA_{1c} Hemoglobin A_{1c} is considered the principal outcome measure for blood glucose control in diabetes mellitus. It was found that the level of HbA_{1c} in diabetic patients is linearly correlated with the peak or area under the curve of the glucose tolerance test.[26] The utility of HbA_{1c} concentration as a monitor of the degree of control in diabetic patients was proven by several studies.[27] They indicated that the concentration of HbA_{1c} in the blood accurately reflects the carbohydrate status of the patient over the preceding weeks to months, and in effect gives a value that integrates the blood sugar concentration over many weeks. A continuous-valued percentage, HbA_{1c} has a normal range of 4.9% - 6.7%, with values below 7% indicative of diabetes patients who have their blood glucose levels under tight control. HbA_{1c} will be measured at baseline during participants' first in-clinic visit (T0) and at three month follow-up.

Patient Satisfaction. This will be assessed via a self-report instrument with the following questions, asking users to place an X on a calibrated line which ranges from 0 to 100% satisfied[28]: (1) How satisfied were you with the information about diabetes that you received today? (2) How satisfied *were you* with the type of help you actually received today? (3) Overall, how satisfied were you with today's visit?

D.3.1.3 Potential Confounding Variables

Functional health literacy. This will be assessed using the s-TOFHLA, a 36-item timed reading comprehension test that uses the modified Cloze procedure;[29] every fifth to seventh word in a passage is omitted, and 4 multiple-choice options are provided. Scores from this self-administered questionnaire (range 0-36), are conventionally categorized into "inadequate" (0-16), "marginal" (17-22), or "adequate" (23-36).[30] This will be measured during the participants' first clinic visit (T0).

Sociodemographic information. Sociodemographic information will include age, gender, marital status, ethnicity, education, and insurance status (previously found to have a significant effect on glycemic control[12]). This information will be collected at baseline.

Health Status. General health status will be measured at baseline using the SF-12.[31]

Cognitive Status. Cognitive status will be assessed at T0 using the Mini Mental Status exam.

Personality. Subject personality will be assessed along the introversion/extroversion dimension (as in the REA experiment described in Section C.1) using relevant scales from Wiggins at T0.[32]

Attitude Towards Computers. Attitude towards computers will be assessed using the Computer Attitude Scales (CAS)[33] at T0. These scales assess users' computer confidence, liking, and anxiety, and the extent to which they perceive computers are useful.

Attitude Towards the ECA. Attitude towards the animated character will be assessed at T1 using the bond subscale of the Working Alliance Inventory (WAI),[34] as modified for the FitTrack study (see Section C.4). The WAI measures the trust and belief that a helper and helpee have in each other as team-members in achieving a desired outcome, and the bond subscale (12 Likert items) assesses the degree to which the helper and helpee like and trust each other.

Diabetes Treatment. Treatment will be defined as diet alone, oral hypoglycemic agent alone, insulin alone, or insulin and a hypoglycemic agent. This information will be obtained from study subjects at baseline (T0) and at the three month follow-up (T2), and will be confirmed by review of electronic medical records. Differences in treatment regimen were previously found to be associated with glycemic control.[12]

D.3.2 Study Population

D.3.2.1 Study Setting: The Geriatric Ambulatory Practice

Bohemia Medical Center (BMC) is a private, not-for-profit, 547 licensed bed, full-service academic medical center affiliated with Bohemia University School of Medicine (BUSM).

The Geriatric Aardvark Practice (GAP) at BMC provides primary care to patients 60 years of age and older who are able to use available transportation to attend the clinic. Currently over 1,100 patients are cared for by the practice; there were 3,300 GAP visits in 2001. Many patients have multiple comorbidities, including cardiovascular disease, renal failure, musculoskeletal disorders, chronic obstructive pulmonary disease, and dementia.

The practice cares for 350 patients with diabetes, 61% of whom are African American (10% are Hispanic and 20% are White), 64% of whom are women, 43% of whom are 75 years of age or older, and 28% of whom are non-English speaking. Medicare is the primary insurer for 85% and 20% are dually eligible for Medicare and Medicaid. About 70% of patients receive home care services and 15% participate in Adult Day Health Programs.

We expect the functional health literacy breakdown in the GAP population to be similar to that found in a study of the health literacy of patients at urban hospitals in Atlanta, GA, and Torrence, CA.[3] This study found that among patients over 60 years of age, 47.9% to 80.5% (depending on site and native language) had inadequate functional health literacy.

D.3.2.3 Eligibility and Exclusion Criteria

Eligibility criteria for this study include:

1. A patient of the GAP clinic.
2. Age 60 years or greater,
3. Have Type 2 diabetes mellitus, with or without complications (ICD-9 codes 250.00-250.90),
4. Understand spoken English, and
5. Score below 80% on a test of general diabetes knowledge as measured by *the Diabetes Knowledge Scales*.

Exclusionary criteria include:

1. Patients with significant cognitive disability (scoring below 21 on the Mini Mental Status exam or with a diagnosis of Alzheimer's disease or other form of dementia, schizophrenia or other psychotic disorder),
2. Patients with significant mechanical functional disability (as determined by the ability to use a touch-tone phone, to ensure they can use the touch-screen computer), and
3. Patients with corrected vision of 20/50 or worse (visual acuity tested with a pocket vision screener, Rosenbaum, Graham-Field Surgical Co, Inc).
4. Patients who have participated in any previous study under this Research Plan.

D.3.3 Sample Size and Power Considerations

We view these experiments as pilot studies, which we anticipate will lead to a larger-scale study of an intervention developed through this work. Given the pilot nature of the proposal, sample size is adequate to detect moderate to large treatment effects. We anticipate enrolling 30 subjects into each arm of each study. Data immediately following intervention will be obtained during the initial study visit and so we do not anticipate any loss to follow-up for these analyses. We conservatively estimate 90% follow-up at 3 months.

Given the proposed sample size, we have 80% power (at the two-tailed alpha 0.05 level) of detecting an effect size (difference in means divided by the standard deviation) of 0.74 immediately following treatment, and 0.78 at the 3 month follow-up. These are considered large effects in the language of Cohen.[35] To estimate the anticipated effect of our intervention, we used data from a study by Dunn, et al, of a diabetes education program that was similar in content to our proposed intervention.[25] This study showed an increase in mean knowledge from before to after an education intervention from 7.61 (sd 3.91) to 11.27 (2.72). Assuming no mean change in the control group, and conservatively assuming a correlation of 0.50 between the baseline and T1 scores, this corresponds to an effect size of 1.08, which is detectible with 98% power given the proposed sample. Given the expected pre-intervention mean for our instrument of 60 (with a sd of 19), the detectable effect size of 0.74 gives 80% power of detecting a 14 points increase in mean knowledge, on average in the intervention group, given no mean change in the controls.

D.3.4 Recruitment and Data Collection Procedures

D.3.4.1 Study Subjects

There are approximately 350 GAP patients, ages 60 and greater, with diabetes mellitus, thereby satisfying eligibility criterion #1-3. Of these, we estimate 28% (98) do not speak English, resulting in 252 patients who also satisfy eligibility criteria #4. A recent study of general diabetes knowledge by patients with diabetes found that patients who had not received diabetes education scored 60 $\pm$19%,[36] indicating that 84% of participants should score below the 80% cutoff, resulting in 212 subjects who satisfy eligibility criteria #1-5. We estimate that 20% of GAP diabetes patients have either significant cognitive disability, significant mechanical functional disability, or visual acuity of 20/50 or worse (exclusion criteria #1-3). Based on previous studies in the GAP, we estimate that 40% of the eligible patients will refuse to participate, resulting in a pool of 102 participants that can be enrolled in the study. Of these, we anticipate that less than 10% will drop out prior to follow-up, leaving approximately 92 study subjects. (We will minimize drop-out by performing the 3 month follow-up visit in their home, and by compensating all study subjects for both visits). This is sufficient numbers to perform the pilot studies, each of which requires 60 study subjects, except study #2 which requires 90.

D.3.4.2 Recruitment and Initial Telephone Interview

Potential participants will be identified from the GAP database by GAP staff.. All patients with Type 2 diabetes who are 60 years of age or older and are cared for at GAP will be sent a letter of introduction and invitation to participate in the study. This letter will describe the study in detail and provide a telephone number for questions about participation. This introductory letter will be signed by the patient's primary care provider. It will tell them to contact their provider or return a prepaid enclosed postcard if they do not wish to be contacted by the study staff. They will also be assured that a decision to refuse to participate would in no way influence the medical care they receive. The letter will describe the study in detail, invite participation and provide a phone number for any questions. Recipients will be informed that they will receive a phone call from a staff person to solicit their participation and to interview them if they decide to participate. These phone calls will be done one week after the initial mailing.

The purposes of the telephone interview are 1) to explain the study to potential participants and answer questions, 2) to obtain verbal informed consent from those who wish to participate, 3) to obtain background socio-demographic and health status information from all persons called to enable participant non-participant comparisons, 4) to partially screen potential participants for eligibility, and 5) to schedule the T0 study visit to coincide with the time of their next clinic visit.

D.3.4.3 Initial Clinic Visit (T0, T1)

At the end of a potentially eligible patient's next visit to the clinic, they will be administered: 1) the Mini Mental Status exam; 2) visual acuity test; and 3) one *of the Diabetes Knowledge* Scales by the project research assistant. If they are still eligible, the research assistant will review the consent and HIPAA

documents and obtain their approval and signature. Next, they are then given the s-TOFHLA health literacy assessment and given the diabetes education pamphlet, at which time they are asked to review the material in the pamphlet and given 15 minutes to do this. Subjects are then randomized into one of the arms of the study. We will use a block randomization protocol to ensure that the number of subjects assigned to each intervention group will balance after every 4 subjects. We will first stratify subjects by health literacy level (inadequate, marginal, or adequate *as defined in D.3.1.3*), then perform separate block randomizations for each stratified group. A block size of 4 is used because of the small number of subjects and the stratified assignment. If the participant has not had HbA_{1c} tested within the prior week, then a blood sample for HbA_{1c} will be taken at this time.

Subjects assigned to the standard-of-care control (CONTROL) condition are then given a second knowledge test, compensated ($25) and discharged. Subjects assigned to the baseline ECA condition (BASELINE-AGENT) interact with the ECA for 15-30 minutes, immediately after which they are given the second knowledge test, compensated ($25) and discharged.

D.3.4.4 Follow-up Clinic Visit (T2)

All participants return to the clinic three months after their initial in-clinic visit for a final diabetes knowledge test. If the participant has not had HbA_{1c} tested within the prior week, then a blood sample will be taken at this time. If participants are not able to come to the clinic, a researcher will visit them at home to administer the follow-up tests and draw the blood sample. All participants will be compensated $25 at the end of the follow-up visit, plus an additional $10 to cover travel costs if they come into the clinic for follow-up.

D.3.5 Analysis

D.3.5.1 The Analysis Plan

The analysis will compare two randomized groups on data collected at study entry, immediately following the educational intervention, and at 3 months follow-up. We anticipate that immediate effects and distal effects may differ, and so our primary analyses will involve separate analyses at these two time points. Preliminary analyses will compare study groups on baseline demographics, health status, co-morbidities, and functional health literacy level. While significant differences are not expected between study groups because of randomization, baseline characteristics found to differ between groups will be considered potential confounders and controlled for in later analyses.

Analyses immediately following intervention T1 will focus on the outcome of knowledge acquisition. Change in knowledge from baseline T0 to T1 will be calculated, and average change will be compared between study groups through the two-sample t-test. …

Analyses at 3 months of follow-up T2 will examine change in knowledge from baseline to T2. Average change in knowledge will be compared across the two study groups through the two-sample t-test…

Similar analyses will examine our exploratory analysis that the intervention will lead to greater increase in glycemic control. Change in HbA1c from T0 to T2 will be calculated, and average change will be compared between groups through the two-sample t-test. Finally, we will perform exploratory analyses of the impact of the intervention (with the embodied conversational agent) on patient satisfaction, using a similar analysis strategy, in which satisfaction will be assessed for each of the three satisfaction questions.

D.7 Issues and Potential Problems

There are some issues that need to be addressed given the study design:

1. Given the high percentage of minority patients in the subject population, linguistic and cultural barriers to communication with the agent may be an issue. The ideal solution to this problem would be to develop several versions of the baseline ECA—one for each ethnic minority in the population. While the underlying dialogue architecture supports adaptation to other languages and cultures (by separating communicative form from communicative functions[37]), the labor required to produce multiple versions of the agent is beyond the scope of this proposal (although a good candidate for future work). To address this issue, we will perform extensive focus group testing with individuals representative of the study population to develop and refine a single version of the agent that is the most an acceptable and understandable to all study participants.

2. Although patients with significant cognitive impairments will be screened out, we expect some number of those included in the study to suffer from mild cognitive impairment, which could impact their immediate and

distal knowledge gains. Although we feel there is great promise for an ECA to diagnose this condition during its dialogue with patients and adapt accordingly, this is beyond the scope of the current proposal.

E. HUMAN SUBJECTS

E.1 Population Characteristics and Minority Recruitment

The subjects of this study will be patients of the Geriatric Ambulatory Practice at Bohemia Medical Center (BMC) and the General Internal Medicine Practice at BMC. These sites care for a population of urban, disadvantaged individuals. We expect that 66% of the sample will be women and 84% minorities.

E.2 Participation of Children

Children will not be used in this study, as the central research questions specifically are addressed at older adults (60 years of age or older).

E.3 Participation of Minorities

The subjects of this study are patients of the Geriatric Ambulatory Practice at Bohemia Medical Center (BMC), of which 61% are African American and 10% are Hispanic. In order to ensure high recruitment of minorities (African American and Hispanic), we will follow the strategies outlined in the NIH Handbook on Minority Recruitment.

E.4 Potential Risks and Protection Against Risks

Participation is strictly voluntary. Potential risks are minimal. All subjects will be compensated for their participation in each study visit (a total of 2 visits) and also compensated for travel expense for the second visit. No telephone interview or solicitation will be pursued for any contact if the patient is acutely ill, just recovered from an illness or expresses reluctance due to a current family crisis. Individuals may be re-contacted if the exclusion is temporary.

E.5 Recruitment and Informed Consent

The processes of recruitment and obtaining of informed consent are described in detail in Sections D.3.4.1 and D.3.4.2. Patient subjects will be provided with a brief description of the study and the nature of their participation. Verbal consent is obtained during the initial telephone interview, and written consent is obtained, along with HIPAA authorization, during the initial clinic visit. A copy of the consent document and HIPAA authorization will also be provided to the participant. Participants will be assured that their decision whether or not to participate will have no effect on their relationship with their health providers or on the care they receive from them. If they choose to participate, they will be told that they can stop any time without jeopardizing their relationship with their providers or the medical care they receive from them.

A HIPAA waiver will be required for the identification of potential study subjects using the protocol described in Section D.3.4.2. Following this, patient medical records will only be accessed once the patient has given their consent, and are only accessed to confirm treatment regimen for their diabetes at two time points in the study (see Section D.3.1.3).

E.6 Protection of Confidentiality

Assurance of confidentiality of all information will be made to all participants. Participants will be told that all conversations and all data derived from the study will be held in strict confidentiality. Data will be handled with the same confidentiality accorded to patients' medical records. Specific procedures protecting participant confidentiality will be as follows: 1) ID number only will identify all data, 2) master lists linking participants information with ID number will be numbered consecutively and prepared before data collection (to ensure accurate accounting). These lists will be kept locked, in duplicate, with access only by the PI, 3) all project staff will sign an oath of confidentiality to ensure their understanding of the terms of confidentiality required, 4) sign-out procedures for all removals from data files will be strictly enforced, and 5) all reports and publications will preserve the participants' anonymity.

E.7 Benefits and Risk/Benefit Ratio

The risks to participants are minimal. The participants will be able to contact the investigators and/or their providers for advice on any problems that might arise. The potential benefits of being in the study may be

better diabetes knowledge and possibly better health outcomes in terms of improved glycemic control by diabetics resulting from a better understanding of their self-care treatment regimen. Each participant will also be contributing to a pilot study, which will answer important questions regarding the feasibility of conversational agents for patient education.

E.8 Data and Safety Monitoring Plan

An independent Monitor, Dr. Howe, in accordance with all NIH guidelines, will carry out a Data and Safety Monitoring Plan. Dr. Howe, one of the researchers in this project, has significant experience in the conduct of studies involving human subjects. He will review the subject recruitment protocol prior to implementation. He will consider whether or not the risks associated with implementation of the protocol are reasonable and are minimized appropriately. Dr. Howe will serve in an advisory capacity to the NIA in order to monitor, review and assess study progress. He will have access to outcome data during the study to ensure that participants are not exposed to unreasonable or unnecessary research risks. Dr. Howe will also review the quality of the data, based on data monitoring reports and other material submitted by the PI. Dr. Howe will assess each study every three months throughout the trial. He will attend research project meetings. He will also forward summary reports to the IRB of Bohemia University School of Medicine and to the NIH. Each summary report will include:

- The date of the review.
- A summary of the adverse events reported.
- A conclusion with respect to progress or need for modification of the protocol.

These summary reports are in addition to all other adverse event reporting procedures required by NIH and the local IRB, and will be distributed to the PI within 30 days after the review.