Available online at www.sciencedirect.com

SCIENCE DIRECT

Interacting with Computers 17 (2005) 711–735

Interacting with Computers

www.elsevier.com/locate/intcom

'It's just like you talk to a friend' relational agents for older adults

Timothy W. Bickmore[a,*], Lisa Caruso[b], Kerri Clough-Gorr[b], Tim Heeren[c]

[a]*Department of Medicine, Medical Information Systems Unit, Boston University School of Medicine, Boston, MA, USA*

[b]*Department of Geriatrics, Boston University School of Medicine, Boston, MA, USA*

[c]*Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA*

Available online 30 September 2005

Abstract

Relational agents—computational artifacts designed to build and maintain long-term social–emotional relationships with users—may provide an effective interface modality for older adults. This is especially true when the agents use simulated face-to-face conversation as the primary communication medium, and for applications in which repeated interactions over long time periods are required, such as in health behavior change. In this article, we discuss the design of a relational agent for older adults that plays the role of an exercise advisor, and report on the results of a longitudinal study involving 21 adults aged 62–84, half of whom interacted with the agent daily for 2 months in their homes and half who served as a standard-of-care control. Results indicate the agent was accepted and liked, and was significantly more efficacious at increasing physical activity (daily steps walked) than the control.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Relational agents; Social interfaces; Human–computer interaction

* Corresponding author. Present address: College of Computer and Information Science, 202 West Village H, Northeastern University, 360 Huntington Avenue, Boston, MA 02115-5000, USA. Tel.: +1 617 638 8170; fax: +1 617 373-5477.

E-mail addresses: bickmore@ccs.neu.edu (T.W. Bickmore), lisa.caruso@bmc.org (L. Caruso), kmclough@bu.edu (K. Clough-Gorr), tch@bu.edu (T. Heeren).

0953-5438/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.intcom.2005.09.002

1. Introduction

Between 2010 and 2030 the number of older adults in the United States is projected to more than double, with the overall fastest growing segment being those age 85 and older (DHHS, 1991). This shift in US demographics towards an older society has motivated a growing body of research in human–computer interaction for older adults over the last decade. While some studies have found that older adults embrace new technology, others have concluded that many older adults have significant problems using standard computer interfaces. The burden of physical and cognitive impairments in this age group may explain some of the difficulties older adults have using computers. Computer anxiety may also be one of many factors that have led to a low adoption rate of computers by older adults. As a group, adults age 55 and over own fewer computers than any other adult age group. Only 25.8% of senior households have a computer compared to 46.0–54.9% for adults aged 25–54. Access to the Internet lags as well: 14.6% of all senior households have Internet access compared to 30.1–35.0% for adults aged 25–54 (NTIA, 1999).

Embodied conversational agents (ECAs)—animated computer characters that simulate face-to-face conversation with users (Cassell et al., 2000)—may constitute an especially effective form of human–computer interface for older adults. One skill that many older adults retain, even with significant cognitive impairment, is the ability to engage in face-to-face conversation. This primal form of interaction is learned in infancy and early childhood and comprises tacit, crystallized knowledge in older adulthood. The ability to engage others is thus available even when individuals have difficulty learning new interfaces that require more adaptive, ‘fluid’ intelligence (Horn, 1982). Face-to-face dialog is multi-modal. It incorporates a wide range of non-verbal and paraverbal behavior to carry semantic content redundantly with speech, enabling individuals with impairments in some channels (e.g. hearing) to compensate by attending to redundant information in other channels (e.g. hand gesture). Face-to-face dialog also has well-established repair mechanisms that enable listeners to request repetition or clarification by a speaker (e.g. communicated by the absence of grounding behavior such as a headnod; Clark and Krych, 2004). Finally, this form of interaction has mechanisms for constraining the interactants’ focus of attention (Grosz and Sidner, 1986). This focus is important for older adults who have difficulty in dividing their attention or dealing with distractions common in current windows-based interfaces (Hartley, 1992).

In addition to task-oriented conversations that ECAs can hold with users, we feel that incorporating social, emotional and relational aspects of face-to-face interaction into a computer agent is also crucial for agents designed to interact with older adults, especially over long periods of time. A significant amount of research over the last three decades has demonstrated the importance of social support on health and well-being in the elderly (Berkman and Syme, 1979; Berkman and Glass, 2000). In addition, studies across many disciplines of human helping have demonstrated the significant positive effects of a high-quality relationship between the helper and the helped on therapeutic outcomes (Horvath and Symonds, 1991). Assuming these effects hold for human–computer relationships, an ideal interface agent for older adults in helping tasks should use relationship-building and maintenance behaviors in addition to being a competent conversationalist, which comprises a kind of interface we refer to as a ‘relational agent’.

Such interfaces move beyond the paradigm of computer as tool, to make systems that are delightful and engaging so that users want to use them over and over again, and that they will trust and really care about, as well as feel cared for by them (Cassell and Bickmore, 2003; Bickmore and Picard, 2004). Although many studies of social and embodied agents have shown these effects in brief interactions (Dehn and Mulken, 2000), an important question raised by this research is whether these results are durable or not; that is, they may simply be the result of a novelty effect that quickly wears off.

In our work, we have developed a relational agent that plays the role of an exercise advisor, designed to be used by older adults in their homes on a daily basis over an extended period of time. One reason we selected this application domain is that the daily protocol—driven by the current government recommendations for daily physical activity—provides ample contact time to establish a social relationship with users over the duration of the study. In addition, older adults are in particular in need of physical activity interventions: only 12% of adults over 75 get the minimum level of physical activity currently recommended by the Centers for Disease Control and Prevention, and 65% report no leisure time physical activity (DHHS, 2000).

In this article, we begin by defining and describing relational agents, arguing that they are uniquely suited to long-term helping applications for older adults, and then present a review of related work. We then discuss the design details of our relational agent application, including both general principles of designing social interface agents for long-term use and issues specific to our application. We then describe our evaluation study and quantitative results, followed by an extended discussion of the ramifications of the study, supported by qualitative results from semi-structured interviews with study participants. We conclude with a discussion of promising future directions of research in this area. In an earlier conference paper we presented a synopsis of the acceptance and usability results from this study (Bickmore, 2005); in the current article we present all results from the study, in addition to a detailed design rationale and discussion of implications and future work.

2. Relational agents

Relational agents are computational artifacts designed to build and maintain long-term social–emotional relationships with users (Bickmore, 2003). These can be purely software humanoid animated agents—as developed in this work—but they can also be non-humanoid or embodied in various physical forms including robots, wearables and hand-held computers. Central to the notion of relationship is that it is a persistent construct, spanning multiple interactions; thus, relational agents are explicitly designed to remember past history and manage future expectations in their interactions with users. Finally, relationships are fundamentally social and emotional, and detailed knowledge of human social psychology must be incorporated into these agents if they are to effectively build relationships with users in the most natural manner possible.

This work uses a relational agent that is a specialized kind of ECA. In addition to the reasons cited in Section 1, an ECA was selected because language is the primary modality used to build human relationships (Duck, 1995), face-to-face conversation is the primary site of human language use, and many of the relational strategies that humans use within conversation are non-verbal (Andersen and Guerrero, 1998).

Although some forms of relationship can be constructed without language (e.g. ‘relationships’ with pets) language is crucial for the development and management of human relationships (Duck, 1994). Given this, our agent uses natural language and synchronized non-verbal conversational behavior to communicate with users. User input, however, is more problematic. While we would like to use unconstrained speech input, we find the state of the art in speech recognition (especially for older adults) and natural language understanding (especially for unconstrained social dialog) does not support this. Consequently, we offer users on-screen, multiple-choice inputs, dynamically updated during each turn of the conversation. While this can be very constraining for some users, properly constructed agent dialog and user response choices can still maintain the feel of social interaction, and was shown in the MIT FitTrack study (Section 3.4) to be sufficient for establishing social bonds between a relational agent and user over time. In addition, this form of user input, which relies on recognition rather than recall and synthesis, may be easier to use for many older adults who suffer from language production impairments (Ulatowski et al., 1985).

3. Related work

3.1. Dialog systems for health communication with older adults

A number of dialog systems have been developed for older adults, particularly in the area of health communication.

The TLC-ACT system was an automated telephone system that used recorded speech output and touch-tone input to promote exercise adoption (Jarvis et al., 1997). Weekly calls were made over a 3-month period to 41 older adults (average age 66), while a ‘usual medical care’ control group received printed information on the benefits of walking. Subjects in the intervention group who walked 15 min or less at baseline significantly increased the number of minutes walked at the end of the intervention compared to controls. In addition, subjects were very satisfied with the system, giving it a score of 8.6 (range: 1 = ‘very dissatisfied’ and 10 = ‘very satisfied’), and 74% of the intervention group rated their satisfaction as 10 of 10.

A number of automated reminder systems have been developed for older adults with mild to moderate cognitive impairment. Perhaps, the most sophisticated system developed in this area is Autominder, which incorporates sensing, planning and scheduling capabilities to adapt to the activities of the older adult user (Pollack et al., 2003). Autominder also communicates to its users in natural language, and uses a custom text generation algorithm to generate reminders that are as effective and non-repetitive as possible. However, it does not engage users in an interactive dialog or have any social aspects to its planned communication.

3.2. Embodied conversational agents for older adults

ECAs have rarely been used as interfaces for older adults. Besides the animated pill used in the Medication Advisor, the only ECA system in this area that we are aware of is the GrandChair system (Fig. 1) (Smith, 2000). In this system, the agent had the appearance of a 6-year-old child, and was designed to elicit narratives from older adult users. The system used user posture and speech intonation to provide continuation prompts at appropriate intervals ('tell me more'). In a study involving 15 grandmothers (aged 55–65) the agent was found to elicit significantly more and longer stories than a text-prompt control condition. Fourteen of the 15 users enjoyed the experience, and all users automatically adopted a child-directed speaking style appropriate for a 6-year-old child.

3.3. Emotional care-taking robots for older adults

Several recent efforts in academia and industry have focused on creating robotic pets for older adults in order to achieve the same beneficial effects found in animal-assisted therapy, namely to decrease stress, anxiety and loneliness and improve mood (e.g. Banks and Banks, 2002). The 'mental commit' robots take the form of cute stuffed animals such as a cat or a harp seal pup, and are designed to foster an attachment with users. One study compared the effects of these robots on older adults in a nursing home with the effects of an identical robot that had a much simpler behavioral repertoire, however, no significant differences were found (Wada et al., 2003). Although the robot was used in group sessions by the same individuals 4 days per week for 3 weeks, and has long-term memory and a reinforcement learning mechanism, it is not possible for it to model a relationship with any

Fig. 1. GrandChair, an early conversational system for older adults.

particular user given that it does not have the ability to discriminate between users. Another study compared the use of a Sony AIBO robotic dog with a stuffed toy dog and a 'clothed' AIBO by a group of older adults with severe dementia. This study found that patients actually interacted more with the stuffed toy than either of the AIBOs, but the differences were not significant (Tamura et al., 2004).

There is also an emerging commercial market for robotic dolls targeted at the older adult market, particularly in Japan. Bandai launched the Primopuel doll in 1999, which is designed to resemble a 5-year-old boy who continually asks to be hugged and entertained. Dream Supply released the Snuggling Ifbot in 2004, intended to be a speech-based conversational partner for the elderly. Tomy recently announced the release of the Yumel doll, which also converses with older adults using speech, and assists users in maintaining healthy behaviors such as good sleep hygiene (AFP, 2005). None of these systems have been formally evaluated.

3.4. MIT FitTrack

The FitTrack system was developed to investigate the ability of relational agents to establish and maintain long-term, social–emotional relationships with users, and to determine if these relationships could be used to increase the efficacy of health communication and health behavior change programs delivered by the agent (Bickmore et al., to appear; Bickmore and Picard, to appear). The system was designed using a client–server architecture, with the client application running on users' home computers. The client bundled an ECA, web browser, input screens, and speech synthesizer (see Fig. 2), while the server maintained the relational database, web server (for forms, educational content, and self-monitoring graphs) and dialog engine.

In FitTrack, interaction dialogs are scripted using a custom scripting language that compiles into Augmented Transition Networks (ATNs; Woods, 1986) so that common sub-dialogs could be factored out and re-used across interactions. Each state in the network generally specifies: (1) an output utterance for the agent that can be tailored at runtime using template-filling (agent non-verbal behavior is automatically generated at system compile time using BEAT; Cassell et al., 2001); (2) a set of allowable user utterances that will be presented to the user as a multiple choice selection menu; and (3) next state specification and actions to perform for each user utterance.

An evaluation of the physical activity advisor agent was conducted in a longitudinal, between-subjects design study, involving 101 MIT students and staff interacting with the agent daily for 30 days. The user–agent relationship was evaluated using the Working Alliance Inventory, which is a standardized measure used in psychotherapy to evaluate the quality of therapist–client working relationships (Horvath and Greenberg, 1989).

The study had three treatments: RELATIONAL, NON-RELATIONAL, and a baseline CONTROL condition. In each study arm, subjects received standard behavioral interventions, including goal setting and self-monitoring via progress charts showing their activity levels over time. All subjects were also provided with daily web pages of educational content on the topic of walking for exercise. All subjects in RELATIONAL and NON-RELATIONAL conditions had a daily 'conversation' with the virtual exercise advisor named Laura about their progress, any obstacles they had to exercising, and

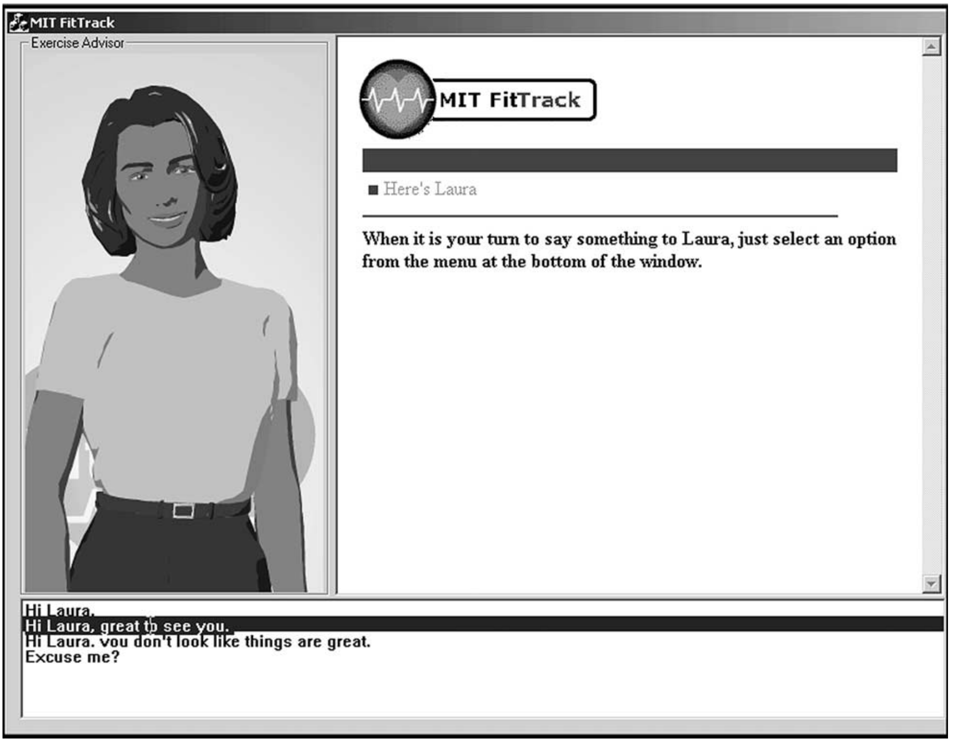

Fig. 2. MIT FitTrack client interface, used in an earlier study.

the educational content. In the RELATIONAL group, the agent also used relational strategies in an attempt to build a working alliance relationship with subjects, whereas in the NON-RELATIONAL condition the relational functionality had been removed.

Subjects in the RELATIONAL condition had significantly higher Working Alliance scores (on the bond subscale of the measure) compared with subjects in the NON-RELATIONAL condition, both at 1 week into the intervention ($t(58)=1.75$, $p<0.05$) and at the end of the 30-day intervention ($t(57)=2.26$, $p<0.05$). Subjects in the RELATIONAL and NON-RELATIONAL groups combined increased their number of days per week over the 30 min per day of moderate-or-greater physical activity goal (during the intervention period) significantly more than subjects in the CONTROL condition, $t(86)=1.98$, $p<0.05$. However, there were no significant differences between the RELATIONAL and NON-RELATIONAL groups with respect to gains in physical activity.

4. Design of a relational agent for older adults

In this section, we first describe some general design principles for relational agents, and then discuss the specific system we implemented for older adult users.

4.1. Principles of relational agent design

Relational agents that will be used for extended periods of time require special design considerations compared to systems that are either only used for brief interactions or do not engage the user in social interaction. An overriding goal in designing these systems is to make them continually engaging to use so that people want to keep using them over and over again.

4.1.1. Model of user–agent relationship

If the user–agent relationship is expected to change over the projected period of use, then a key element of the system is a model of this relationship, methods for updating it, and its use in planning dialog and other interaction behaviors. There are many kinds of relational models (see Bickmore and Picard, to appear for a summary), but perhaps the simplest is one that varies along a fixed trajectory based on number of interactions or total contact time with the user. For example, an agent might be designed so that initial interactions with it by a particular user will be relationally distant and professional, but gradually become more personal and social over time, as reflected by increasing turns of social dialog and increasing frequency of other kinds of relational behavior. This is the relational model used in the FitTrack system (for a more sophisticated example, see Cassell and Bickmore, 2003).

4.1.2. Use of relational behavior

There is a wide range of relational behavior an agent can use in attempts to manipulate or maintain its relationship with a user, and, conversely, that the agent could recognize in the user's behavior to infer their relational stance. These include social dialog, empathy, meta-relational communication (talk about the relationship), humor, and self-disclosure; see Bickmore and Picard, to appear for an extended discussion). Empathy, in particular, is frequently mentioned in the helping literature as the most important factor in establishing a working alliance (Gelso and Hayes, 1998). Empathy may also be especially important in systems developed for older adults, since older adults are much more prone to relating painful information about their lives in social dialog than younger adults (Coupland et al., 1991).

4.1.3. Persistent memory

A crucial element in a relational agent system is a memory of past interactions with a user. While this is theoretically not a necessary element of the relational model (relationships can be characterized categorically or via specific expectations about the future; Bickmore, 2003), it is necessary for the proper use of many relational behaviors, such as talking about the past and future together (Planalp and Benson, 1992; Planalp, 1993) and continuity behaviors (talk about the time spent apart; Gilbertson et al., 1998). Several subjects in the MIT FitTrack study reported that they first noticed that there was something special about the agent when it remembered something about them from a past interaction.

Persistent memory should ultimately be represented as an episodic store recording all details of all (or key) past interactions with the user. At a minimum, however, it can be designed to record specific facts that can be referenced in future conversations. Examples in the physical activity domain include remembering the name of a user's walking buddy or favorite walking location, as well as purely social (off-task) facts, such as the user's favorite television program and whether they had any big plans for the upcoming weekend or not. FitTrack uses a slightly more sophisticated version of persistent memory, in which it records facts indexed to specific conversations, so that the agent can tell, for example, whether the user has not been feeling well for the last four interactions.

4.1.4. Variability in agent behavior

People continuously vary their language in subtle ways in response to changes in context, mood, interaction history, and, of course, relationship with their listener. Imagine how strange it would be if every time you ran into an acquaintance they used *exactly* the same language in their conversation with you. Not only would this seem unnatural, but it would also send the message that their past history with you is unimportant. Thus, in order to establish a social bond with the user and to maintain their engagement over long periods of time, some variability in a relational agent's verbal and non-verbal behavior is important. In the MIT FitTrack study, many subjects felt that the agent was repetitive after the first week or two. This was not just an annoyance; several subjects said that it had a significant negative impact on their engagement with the system and on their motivation to exercise.

Determining appropriate types and amounts of behavior variability to maintain engagement remains an important area of research. There are several approaches to providing variability, however, that are used in FitTrack and are relatively straightforward to implement. First, at each state in the dialog, multiple agent utterances are specified and randomly selected from at runtime, providing some variability in the agent's verbal and non-verbal behavior for every possible utterance it might make. Second, the use of items from persistent memory to both condition logical branches in the dialog and to fill in utterance template slots provides variability over time. Overall dialog structure is also varied, as described in Section 4.1.5.

4.1.5. Dialog structure

Some variability in overall dialog structure is important as well. However, if the ATN-scripting approach to dialog is used, this can be an imposing task, especially if each interaction is scripted as a separate dialog and every dialog has a significant amount of branching. This was the approach taken in the MIT FitTrack system, and it clearly is not scalable. In the system for older adults, we use a fixed top-level dialog structure, but with 'hooks' at each location at which dialog variants can be inserted (e.g. for social dialog). This is coupled with a database of dialog fragments that are indexed by interaction so that, rather than having to write entire top-level scripts for each interaction, only the elements that are different from the standard task-oriented interaction need to be authored.

The dialog structure used in the system for older adults is shown in Fig. 3. Social dialog fragments can be inserted between any of the steps in the above structure, based on the number of interactions the agent has had with the user. In addition, many of these fragments are written so that they could be re-used, such as talking to users about the weather, their weekend plans, or whether they were enjoying their exercise or not.

1. Greeting
2. Ask about user's physical & emotional state ("how are you?")
3. Talk about events that have occurred in user's life since previous contact.
4. Ask about user's pedometer readings (includes showing self-monitoring graph of steps walked vs. agent-suggested goals over time)
5. Review user's exercise progress with respect to long-term goals and short-term commitments (includes positive reinforcement and problem solving)
6. Negotiate next walking commitment (includes tips and affirmations)
7. Ask about when user will contact the agent again
8. Farewell
9. Introduce educational content page for the day

Fig. 3. Top-level dialog structure for new FitTrack system.

4.2. *FitTrack for older adults*

FitTrack was adapted for a population of older adult users (aged 65 and over) with little or no previous computer experience. The system was designed to be easy to use, with a very consistent and intuitive user interface. To achieve these objectives desktop computers were chosen, but instead of using a keyboard and mouse, large (17″) color touch screen monitors were used. The ECA display area was enlarged to fill most of the screen in order to accommodate visual impairments, and the scrolling list of user input options at the bottom of the screen was replaced with a list of large buttons with enlarged text along the right edge of the screen (see Fig. 4). The self-monitoring graph and educational content was displayed by temporarily replacing the ECA, in order to maintain large font sizes throughout (Fig. 5, left and center, respectively). Subjects were required to enter a daily

Fig. 4. Interface for new FitTrack system, showing conversational agent.

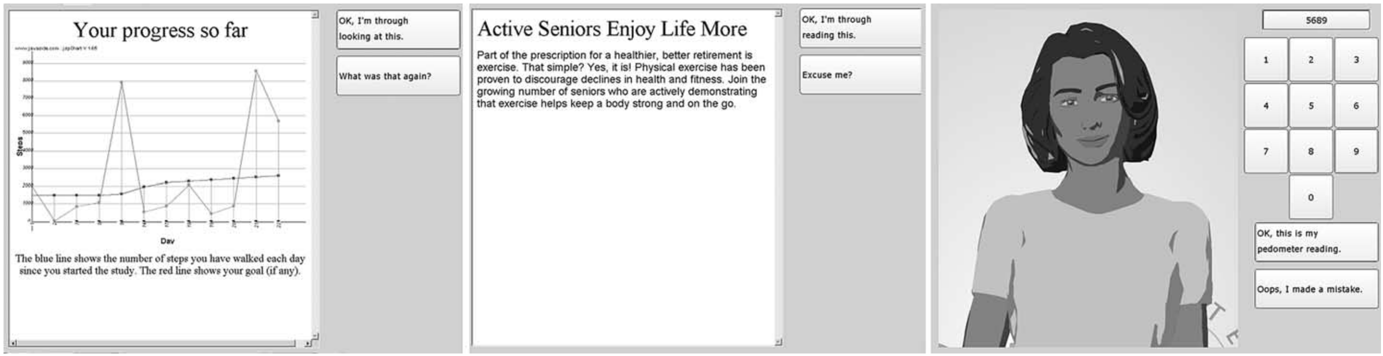

Fig. 5. Additional interface displays for new FitTrack system: (left) self-monitoring chart; (middle) educational page; (right) pedometer entry.

pedometer reading. For this, a numeric keypad entry screen was used so that it appeared in the same area as all other user inputs (Fig. 5, right).

The FitTrack architecture was redesigned to run stand-alone on one computer to eliminate the need for Internet access, with a local database maintaining the persistent memory and logs of all user interface actions for evaluation and troubleshooting. Participants were provided with a dedicated-use PC and table for use during the study period. The system was configured so that participants simply pushed the PC power button to automatically run the relational agent interface, conduct the 5–10 min daily interaction, and automatically shut down the system once completed. In order to prevent loss of data due to system crashes or theft, a concern given the implementation neighborhoods, the system was designed to walk the user through a weekly floppy-disk backup procedure.

Most of the dialog content from the original MIT FitTrack system was re-authored for an older adult population. The conversational scripts used by the agent were designed to be used daily over a 2-month period of time, following the structure outlined in Section 4.1.5. The daily conversation involved having participants enter the number of steps they had walked on the previous day, showing them a self-monitoring graph of their progress, providing positive feedback when warranted, discussing obstacles to exercise, and negotiating a walking goal for the following day which was ‘shaped’ to gradually allow them to work from their baseline level up to their 2-month goal. Since there was very wide variability expected in the amount of walking our participants would be performing given

1. A: Hello, Bertha. Great to see you.
2. U: Hi Laura.
3. A: How are you feeling today?
4. U: I'm not feeling so great.
5. A: I am sorry to hear that.
6. A: Tell me more about how you are feeling.
7. U: I'm tired.
8. A: How tired are you feeling?
9. U: I'm pretty beat.
10. A: Good grief, I hope you’re not getting sick. Can you get some rest soon?
11. U: I think so.
12. A: That is great. I hope you come back feeling refreshed.
13. U: Me too!
14. A: So, are you from the East Coast originally?

…

24. A: So, is there anything I can do to help make this a more productive working relationship?
25. U: Nope, everything is OK.
26. A: That's wonderful.
27. U: Yeah.
28. A: So, let us talk about your pedometer readings.

…

40. A: Looks like mission accomplished on the exercise.
41. A: We're doing some great work together.

…

Fig. 6. Partial transcript of sample conversation between subject and agent (A, Agent; U, User).

the heterogeneity of physical function among older adults, we set different 2 months goals for participants based on their first week's baseline data. The goal was: 5000 steps/day if their baseline was below 5000 steps/day; 10,000 steps/day if their baseline was 5000–10,000 steps/day, otherwise it was set to maintain their baseline level. In addition, the agent used the relational behaviors used in the MIT FitTrack system, including social dialogue, empathy exchanges, meta-relational communication, and non-verbal immediacy behavior, adapted for the new subject population. For example, social dialog about MIT and Cambridge was replaced with dialog about television shows and Boston neighborhoods. Fig. 6 shows part of a transcript from one of 211 interactions conducted with subjects during the study.

The 60 pages of educational content were assembled from publicly available web pages on exercise topics, particularly walking for older adults, from the American Council on Exercise (ACE, 2005), the National Institutes of Health, and the Centers for Disease Control and Prevention.

4.3. Pre-test study

Prior to the intervention study, the interface was pre-tested with five participants in the Geriatric Ambulatory Practice (GAP) clinic at Boston Medical Center, the primary safety net hospital in the Boston area. All participants found the system easy to use with the exception of the floppy-disk backup procedure. Two participants accidentally re-booted the PC while attempting to eject the disk.

Four of the five reacted positively to the agent and indicated that they would be interested in having the agent in their homes to talk to them daily about exercise. We also gave participants 11 designs for ECA characters developed by an animator for the project, intended to more closely match the subject population in age, race and body build.

Fig. 7. Top-rated animated conversational agent choices by geriatric patients.

Subjects were asked to pick the character they would feel most comfortable working with. Fig. 7 shows the top-rated choices.

Based on these results we eliminated the floppy-disk backup procedure from the protocol (no loss of data actually occurred during the intervention study). Unfortunately, given budget and scheduling constraints, the 'Laura' character used in the MIT study was re-used rather than using the selected character, although we plan to use these characters in future studies. This turned out to not be a problem, since there were no complaints about Laura's appearance from intervention study participants. In fact, the Laura character had been intentionally designed to be racially ambiguous, and one subject in the study commented that they liked the fact that Laura was a 'person of color'.

5. Evaluation of FitTrack for older adults

To evaluate the acceptance and efficacy of the exercise advisor agent by older adults we conducted a randomized trial comparing participants who interacted with the agent daily in their homes for 2 months (RELATIONAL) with a standard of care control group (CONTROL). The CONTROL group used a physical activity intervention that had already been designed and planned for use in the GAP clinic, and involved giving subjects pedometers and printed materials on the benefits of walking for exercise.

The primary hypotheses evaluated in the study were:

H1. Subjects in the RELATIONAL group would enjoy using the relational agent on a regular basis.

H2. Subjects in the RELATIONAL group would perform more physical activity than subjects in the CONTROL group.

We were also interested in whether the use of a relational agent by older adults would lead to increases in well-being and, more specifically, decreases in loneliness. This was motivated by the general association found between social support and well-being in the elderly, and since the use of animal-assisted therapy has been shown to reduce loneliness in older adults (Banks and Banks, 2002).

H3. Subjects in the RELATIONAL group would have significant increases in well-being and decreases in loneliness compared to subjects in the CONTROL group.

The study took place in Boston neighborhoods between July and December, 2003, and was approved by the Institutional Review Board at Boston Medical Center.

5.1. Procedure

Subjects were referred by GAP physicians and nurses if they met the study eligibility requirements—(1) no significant cognitive impairments, (2) English speaking ability, and (3) ability to begin a new exercise program—and if they expressed an interest in participating in the study. Following GAP referral, subjects were contacted by telephone to complete screening and to schedule an initial meeting with an experimenter in their home. During this meeting, consent was obtained, demographic, literacy, health, well-being and loneliness questionnaires were completed, and pedometer use was demonstrated. Subjects were instructed to keep track of their daily steps on a paper log sheet. Subjects in the RELATIONAL group then had the computer installed in their home, after which they participated in a brief training session in which they conducted an initial interaction with the agent under the direction of the experimenter. They were asked to use the system daily, but informed that they could miss a day or two during the 2-month study period. Subjects in the CONTROL group were left with the printed materials on exercise from the GAP.

At the end of 2 months, an experimenter conducted a follow-up meeting with subjects at which well-being and loneliness questionnaires were completed again and the paper log sheets of steps walked were collected by the experimenter. For subjects in the RELATIONAL group, a questionnaire evaluating various aspects of the intervention was also completed as well as a 15–30 minute, tape recorded, semi-structured interview with the experimenter, after which the experimenter dismantled and removed the computer system.

5.2. Subjects

Twenty-one participants were recruited into the study based on referrals from the GAP clinic: 10 in the RELATIONAL group and 11 in the CONTROL group (see Table 1). Participants ranged in age from 63 to 85 (mean 74.0), were 86% female, and 76% African

Table 1
Subject demographics

		CONTROL	RELAT'L	ALL
Number of subjects		11	10	21
Withdrawals from study		2	2	4
Age		74.2	73.8	74.0
Gender	Male	27%	0%	14%
	Female	63%	100%	86%
Lives alone		27%	30%	29%
Body mass index (BMI)[a]		30.4	27.9	29.2
Overall health (SF-12)[b]	Physical (PCS-12)	41.6	45.8	43.4
	Mental (MCS-12)	56.6	50.7	54.1
Ethnicity	African American	64%	90%	76%
	Caucasian	27%	10%	19%
	Other	9%	0%	5%
Computer experience	Never used	27%	50%	38%
	Used a few times	27%	30%	29%
	Use regularly	45%	10%	29%
	Expert user	0%	10%	5%
Literacy score[c]		65.4	61.6	63.6

[a] BMI of 25–29.9 is 'overweight', 30 or more is 'obese' (NHLBI, 1998).
[b] SF-12 means for US citizens: age 65–74, PCS-12 = 46.36, MCS-12 = 55.31, for age 75 or greater, PCS-12 = 38.68, MCS-12 = 53.53 (Ware et al., 1996).
[c] Literacy score below 75 is 'low literacy' (Lobach et al., 2003).

American. Seventeen (77%) were overweight or obese (based on body mass index from reported height and weight), and 19 (86%) were scored as having low reading literacy (Lobach et al., 2003). Eight (38%) reported never having used a computer before and another six (29%) reported having used one 'a few times'.

There were no significant differences between the groups with respect to age, body mass index or overall health. Overall health scores from the SF-12 were near the national averages based on age (Ware et al., 1996). Participants were compensated for their time.

5.3. Measures

Interaction History was recorded in log files on the subjects' computers that kept track of all actions RELATIONAL participants took with their system.

Steps Walked was measured by a pedometer. Participants were instructed to write their steps down each day on a log sheet, and subjects in the RELATIONAL group were also prompted to enter these into the computer during conversations with the agent.

Satisfaction with, Repetitiveness, Friendliness, Informativeness, Interestingness, Liking of, and Trust in the agent (for subjects in the RELATIONAL group) were measured by single items on seven-point semantic differential scales at the end of the intervention, as were *Ease of Use* of the system, *Desire to Continue* using the system, and *Relationship* to the agent.

Well-being was measured using the Satisfaction with Life Scale (Diener et al., 1985) and the UCLA Loneliness scale (Russell et al., 1980), administered to all subjects at the start and end of the intervention.

Overall health was measured using the 12 Item Short Form Health Survey (SF-12), with subscales for physical health (PCS-12) and mental health (MCS-12) (Ware et al., 1996).

In addition, semi-structured interviews were held with each participant during the follow-up meeting at the end of the intervention.

5.4. Results

5.4.1. Usage

Two participants in the RELATIONAL group and two in the CONTROL group withdrew from the study before the end of the 2-month intervention, all reportedly due to health problems with themselves or a family member. In addition, one participant was found to have not turned the system on once following the intake meeting, even though she claimed to have used it. Following an intent-to-treat protocol, data from all subjects are included in the quantitative results.

For subjects in the RELATIONAL group, actual use during the 60-day intervention ranged from 0 to 54 interactions, with 90% of subjects averaging at least one contact per week, 40% averaging at least two contacts per week and 30% averaging at least three contacts per week. A typical usage pattern was daily during the first week, tapering off to once or twice a week by the end of the study period. When asked if they looked forward to the interactions with the agent, subjects gave a range of responses, with most (75% of those who commented) responding positively:

> Yes, I … yes. Because those two nights I forgot—I think, maybe I had been out late or whatever—but I was really surprised, I was like ‘Oh I forgot Laura.’ Then I’d turn the light on and talk to her. But it was something I looked forward to, I’d say it was my little night cap.
>
> I enjoyed coming in and turning it on, and talking with her.
>
> I can’t say that I looked forward to it. If I did I would have called in every day.

RELATIONAL participants also indicated that they would like to continue using the system, giving this an average rating of 6.4 (range: 1 = ‘not at all’ to 7 = ‘very much’).

5.4.2. Usability

All subjects found the system easy to use, with an average rating of 1.9 on a 1 (‘easy’) to 7 (‘difficult’) scale. Except for some problems entering pedometer steps (described below) and a few other minor problems, none of the subjects reported having any significant difficulty using the system.

A number of participants did have problems using the touch screen keypad to enter their pedometer steps. Compared to their written records they made errors in data entry 49% of the time, often by dropping or duplicating a digit, causing an order-of-magnitude error. This was a major problem if it occurred during the first week when the system was collecting baseline data, as it caused the agent to set an unrealistically high goal for two participants. Two participants also reported problems with touch screen alignment that made selection difficult.

Participants felt that the simulated conversation worked reasonably well. When asked if they felt that they and the agent understood each other, participants rated this at 5.4 (with 1 = ‘not at all’ and 7 = ‘very much’). Only one instance of a problem talking to Laura was reported: the participant thought the agent asked her “are you tired?” when she was really asking “are you retired?” causing the agent to ask her a series of inappropriate follow-up questions. Several participants mentioned that they could not express themselves completely using the constrained, multiple-choice interaction:

> When she ask me questions … I can’t ask her back the way I want.
>
> I felt that she was programmed to answer. She was programmed to listen to the questions that you put on the screen. She would ask a question and I would have a choice, one, two, three, four. But I could never explain. Or she could never follow up, or follow through with another question.

When asked, participants universally said they would have preferred speaking to Laura, rather than using a touch screen. However, they preferred the touch screen over typing, even though it meant having a much more restrictive conversation:

> The only difference is that even if you wanted to say more it was, you know, you could just touch, you couldn’t type, to say something. That was the difference. But, I believe that I would have handled this better. I’m rarely typing in my words. Because I’m not too familiar with computers.

Several participants mentioned that they enjoyed the social dialog with Laura and would have liked the opportunity to chat more with her:

> She says ‘Good Morning Camille’ [laughs]. Yeah, it’s nice. I liked that. You know a lot of more people would like that because they’re lonely.
>
> I told you I like to talk, so that was, you know, good chit-chat’n. I found it very good, you know, chit-chat’n with her.
>
> I wanted to have more conversation with her.

Participants also mentioned that they would have liked to talk more about their health problems with Laura, especially as they related to their ability to do more walking:

> There was a couple of times when I wasn’t feeling good I didn’t go walking, you know. And I felt that. So, I would have loved to say to her, ‘well Laura I think … ’ and she say to me ‘I hope you feel better.’ I think that came up. I said ‘I’m getting the flu,’ and I wanted to say to her ‘I wasn’t feeling well’ so I won’t go walking.
>
> I don’t think she understood about my condition. No, not me. But, probably someone else. Because she kept saying you have to do 10,000 steps … no way could I do that.

5.4.3. Physical activity

Comparisons between the RELATIONAL and CONTROL groups on daily recorded pedometer steps were based on generalized estimating equations (GEE) regression models for longitudinal data. GEE regression accounts for the repeated-measures nature of the data by modeling the correlation between repeated observations from the same subject, and adjusting both parameter estimates and standard errors for this correlation. GEE regression has several advantages over traditional repeated-measures ANOVA, in that it

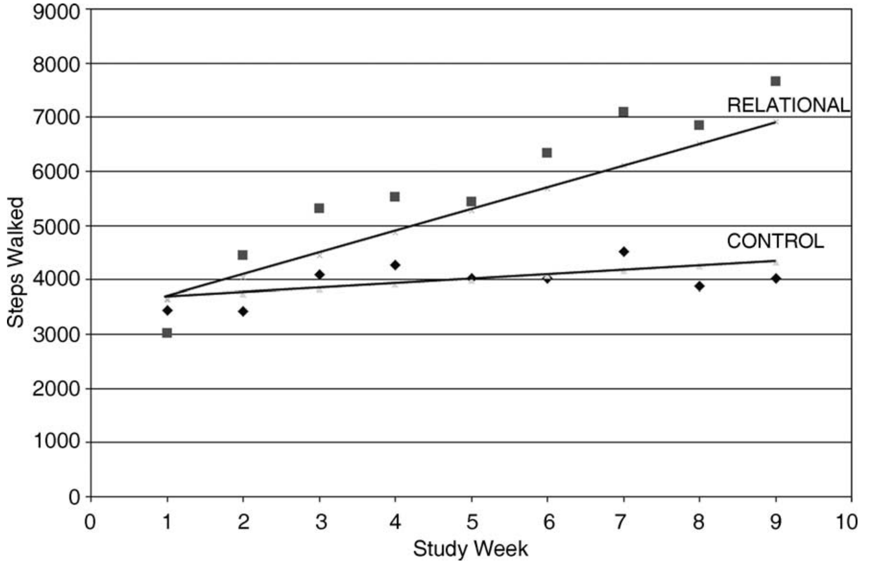

Fig. 8. Study results: subjects who used relational agent walked more.

allows unequally spaced observations over time and allows inclusion of subjects who are missing data from some follow-up points. This approach also allows the flexibility of regression modeling in describing trends over time and controlling for other covariates.

We modeled a linear increase in steps over the study period, and allowed a differential increase in steps in the RELATIONAL vs. CONTROL group by including an interaction term between time and study group in the model. The estimated slope (increase per week in mean steps walked) for the CONTROL group was 83.9, while the slope for the RELATIONAL group was estimated at 411.1 (see Fig. 8). The difference in slopes was significant ($p=0.004$). There was no significant increase in steps over the study period in the CONTROL group ($p=0.295$ for the test of 0 slope), while the increase in steps over time in the RELATIONAL group was significant ($p=0.001$).

5.4.4. Well-being and loneliness

There were no significant changes in the Satisfaction With Life Scale scores for either the RELATIONAL (paired $t(9)=0.34$, $p=0.75$) or CONTROL (paired $t(7)=0.80$, $p=0.45$) groups during the intervention, nor were there significant differences between the two groups on this measure either before ($t(18)=0.557$, $p=0.58$) or after the intervention ($t(17)=0.316$, $p=0.76$).

The CONTROL group did see a significant decrease in loneliness over the intervention period (paired $t(7)=2.74$, $p<0.05$), but this was not reflected in the RELATIONAL group (paired $t(9)=0.50$, $p=0.63$). There were no significant differences between the two groups on this measure either before ($t(18)=0.658$, $p=0.52$) or after the intervention ($t(17)=1.635$, $p=0.12$).

5.4.5. Ratings and perceptions of relational agent

Reviews of the relational agent were mostly positive, as shown in Table 2:

> She's nice. She's really good. Really good. She asks you the right questions. She tells if you if you're not doing up to par, you know, and all that. And if you're doing good, she'll tell you. If you're not she'll tell you. And it's honest. And it works. It really does. I like it. I like talking to her.
>
> She was nice and friendly, and honest, with the stuff she say.
>
> She's nice. She's really good. Really good.

Table 2
Ratings of relational agent

Measure	Min (1)	Max (7)	Mean
Satisfaction with	Not at all	Very	5.4
Liking of	Not at all	Very much	6.3
Trust in	Not at all	Very much	6.4
Relationship with	Stranger	Close friend	6.8
Friendly	Not at all	Very	6.7
Informative	Not at all	Very	6.5
Repetitive	Not at all	Very	4.8
Interesting	Boring	Interesting	6.4

> I enjoyed her, I enjoyed Laura, and I'm quite sure somebody else would.

Eight of the 10 subjects in RELATIONAL anthropomorphized the agent to some degree, putting Laura in a category somewhere between a computer and a person:

> You'd be talking to her and sometimes you forget and think she's a real person.

> I would ask her of course and then I'd say 'oh this is not a person' [laughs]. THIS IS NOT A PERSON [laughing]. You know after you talk to her for so long you get to thinking of Laura as a person.

> Well, just turning on the computer every night, and realizing that this is not a human being, though I am carrying on an intelligent conversation with this, with Laura. And, like I said after maybe the third or fourth night, that thought wasn't even in my mind anymore. I was not talking to a computer. I realized it was not a human being but it was not a computer.

> A telephone is a machine. But, I would not say that Laura is a machine.

> I don't feel that it gets the true feeling, the true experience of a person. Like a real live person. … It was an animation. You know those questions [on the follow-up questionnaire] were inappropriate.

Two subjects compared talking to Laura to talking to a character on TV, or having people on TV tell them what to do:

> It was like talking to the TV. I talk to the TV, too. [laughs]

There were many comments about the agent–subject relationship and relational dynamics. The general relational trajectory seemed to be that the relationship and interaction seemed strange at first, became more familiar over time, with liking and trusting of and even caring for the agent common by the end of the 2 months. Several participants said they also felt that Laura liked and cared about them:

> Um, she took getting used to. But, um, by maybe the third or the fourth night she appeared to be familiar, you know…

> By the way that she sound, she sound like she like me.

> I remember one weekend I went to Wareham… You know, I began to feel bad about Laura, stuck in that box.

> It was funny when it stopped. [strangeness felt after the last interaction]

Rather than serve to further isolate older adult subjects, two subjects mentioned that Laura actually provided a form of social network support for them. One said that talking to Laura became a regular activity that she did with her granddaughter watching, while another mentioned that her friends would often remind her to talk to Laura:

> I brought my friends up here a couple of times to listen to her. My girlfriend she came upstairs with me and I show it. She say 'what's that?' and I say 'let me show you.' So I talk to her. So every time she talk to me she say 'Did you talk to Laura last night?.'

6. Discussion

Overall, the study was successful in demonstrating the acceptance and usability of a relational agent by older adults—even though most of the subjects in the RELATIONAL group had little or no computer experience—and in demonstrating the efficacy of this agent in motivating them to walk more. We conclude that our first two hypotheses regarding enjoyment of the system and health behavior change efficacy were confirmed.

Satisfaction with the overall intervention was very high, with most participants acknowledging that it was for their benefit:

> I appreciated having that kind of a reminder, because I don't have anybody who will tell me what to do, to remind me, you know, to get up, get out and get some fresh air.

> It was the best thing that happened to me, to have something that pushed me out and get me walking.

Participants in our study had very few problems using the PC based system or talking to the agent:

> That is so easy. That is so good. Regular computers I don't do. But, that was so easy, even a baby could do that.

The only notable exception was the part of the interaction that was non-conversational: pedometer step entry.

Both survey results and interview responses indicated that most participants in the RELATIONAL group felt that they had established some kind of social bond with the system over the duration of the intervention and that this bond, at a minimum, served as a motivator for them to continue interacting with the agent.

Our third hypothesis, regarding the impact of the relational agent on well-being and loneliness, was not confirmed. This may be due to the short intervention interval, the small number of subjects, and especially the even smaller number of subjects who lived alone: most subjects either lived with their spouse or in a multi-generational household. The significant decrease in loneliness in the CONTROL group is a little harder to explain, but may be due to the small (but not significant) increase in physical activity experienced in this group and the accompanying increase in social contact that is common with walking for exercise. Why this effect was not seen in the RELATIONAL group is an open question.

7. Future work

There are a number of important and interesting directions for future research on relational agents for older adults. First is that more sophisticated models of dialogue planning should be developed, so that users can engage in richer conversations and can more freely express themselves. Doing this while maintaining the ease of use of the multiple-choice selection input modality will be a very challenging undertaking. Replacing the touch screen with automated speech recognition would have been welcomed by all participants in this study, but commercial systems would need to be thoroughly evaluated to ensure that they could provide high enough reliability given the variability and differences in voice quality in older adults compared to the voice models they were developed for.

Many participants indicated that interactions with Laura became somewhat repetitive over the course of the 2 months, and the solution to this problem remains an interesting area of research. In addition to the techniques described in Section 4.1.4, other possible solutions include: providing a wider range of dialog scripts; more variability in agent utterances for any given dialog state; and the incorporation of external data that is dynamically updated (e.g. weather reports). One subject suggested that just having Laura change her clothing occasionally would have helped break the monotony. Ultimately, what is needed is the use of dialog planning and text generation techniques to dynamically synthesize the agent's behavior and dialog so that it can be subtly varied every interaction, for example, based on simulated agent mood or the user's affective state.

The general relationship between user–agent social bond, system usage and behavioral outcomes needs to be more thoroughly explored. Certainly social bond and engagement influences usage and there is some dose–response association between usage and outcomes, but the details need to be teased out in a much larger study.

Future systems like FitTrack that rely on sensor readings (such as pedometer step counts) should be designed to automate the transfer of data from the sensing devices to the computer to eliminate user data entry errors.

There are also some issues involved with just installing computers in the homes of older adults, especially those with minimal computer expertise. First, many older adults live in very small homes filled with possessions accumulated over a lifetime. Just finding an appropriate space for a desktop computer system can be very challenging. Two subjects had to put the computer on part of their dining room table. Second, multi-generational households with small children can be problematic when installing fragile electronic equipment such as a home computer. One woman dealt with this by including her granddaughter in all of her conversations with Laura. Finally, one participant was concerned that the computer, which was installed as a stand-alone system, was somehow collecting information about her and broadcasting it back to the hospital.

8. Conclusion

In this article, we have examined the use of a new kind of computer interface—relational agents—for older adult users, specifically in the context of a health education and behavior change application. Based on this experience, we believe that the caring, social, anthropomorphic interface enabled the system to be readily accepted, usable with minimal training, and efficacious in achieving a desired behavioral outcome with this user population.

Relational agents may prove to be an especially effective modality for delivering health communication to older adults with low functional health, reading, or computer literacy, problems that are especially acute in underserved populations. In a 1995 study of 2659 predominately indigent and minority patients at two urban public hospitals it was found that 86% of the English speaking elderly had inadequate or marginal functional health literacy, defined as the ability to read and follow basic medical instructions (Williams et al., 1995). A relational agent that uses the universally understood and non-threatening format of face-to-face conversation, relying minimally on text comprehension and using multi-modal cues to maximize comprehension, may be an ideal platform to provide an effective automated health educator with unbounded patience and empathy for these patients.

Relational agent systems such as the one presented here may also be effective in nursing home settings for engaging residents in social conversation and motivating them to engage in social activities, in addition to motivating them to obtain physical activity. Studies have shown that these residents typically spend over half of their time alone doing nothing, leading to loneliness, low self-esteem, depression and, consequently, an overall low quality of life. In addition, because residents spend such a large portion of their day sitting or laying down, they are at great risk for muscle atrophy and skin breakdown (Ice, 2002).

While social and relational interaction is certainly not appropriate for all kinds of computer interfaces, we feel that health communication with older adults is one area in which these interface features are not only efficacious but are sincerely appreciated by users.

Acknowledgements

This work was supported by a grant from the Evans Foundation at the Boston University School of Medicine. Thanks to Ann McDonough, RN, for her help in recruiting subjects from the GAP clinic. Thanks also to Jennifer Smith, Rosalind Picard, Ramesh Farzanfar, Jeffrey Mignault, and Robert Friedman for their many helpful comments on this article.

References

ACE 2005. American Council on Exercise. Retrieved from the World Wide Web onApril 9, http://www.acefitness.org.

AFP 2005. As Japan goes grey, toymakers design dolls for the elderly. AFP Online, 2005, available at http://www.antara.co.id/en/seenws/?id=1109, accessed 18 Aug.

Andersen, P., Guerrero, L., 1998. The bright side of relational communication: interpersonal warmth as a social emotion. In: Andersen, P., Guerrero, L. (Eds.), Handbook of Communication and Emotion. Academic Press, New York, pp. 303–329.

Banks, M., Banks, W., 2002. The effects of animal-assisted therapy on loneliness in an elderly population in long-term care facilities. Journal Gerontology Medicine Sciences 57A, M428–M432.

Berkman, L., Syme, S., 1979. Social networks, host resistance, and mortality: a nine year followup study of alameda county residents. American Journal of Epidemiology 109, 186–204.

Berkman, L., Glass, T., 2000. Social integration, social networks, social support, and health. In: Berkman, L., Kawachi, I. (Eds.), Social Epidemiology. Oxford Press, New York.

Bickmore, T., 2003. *Relational agents: effecting change through human-computer relationships*, Massachusetts Institute of Technology PhD Dissertation, Cambridge, MA: MIT.

Bickmore, T., 2005. Acceptance and Usability of a Relational Agent Interface by Urban Older Adults. *Proceedings of CHI'05*, Portland, OR.

Bickmore, T., Gruber, A., Picard, R. (to appear). Establishing the Computer-Patient Working Alliance in Automated Health Behavior Change Interventions. *Patient Education and Counseling*.

Bickmore, T., Picard, R., 2004. Towards Caring Machines. *Proceedings of CHI'04*, Vienna.

Bickmore, T., Picard, R., (to appear). Establishing and Maintaining Long-Term Human-Computer Relationships, *ACM Transactions on Computer Human Interaction*.

Cassell, J., Bickmore, T., 2003. Negotiated Collusion: Modeling Social Language and its Relationship Effects in Intellient Agents. User Modeling and Adaptive Interfaces 13, 89–132.

Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), 2000. Embodied Conversational Agents. The MIT Press, Cambridge, MA.

Cassell, J., Vilhjálmsson, H., Bickmore, T., 2001. BEAT: *The Behavior Expression Animation Toolkit. Proceedings of SIGGRAPH '01*, Los Angeles, CA.

Clark, H., Krych, M., 2004. Speaking while monitoring addressees for understanding. Journal of Memory and Language 50, 62–81.

Coupland, N., Coupland, J., Giles, H., 1991. Language, society, and the elderly. Blackwell, Oxford.

Dehn, D., Mulken, S.v., 2000. The impact of animated interface agents: a review of empirical research. International Journal of Human-Computer Studies 52, 1–22.

DHHS, 1991.Anon., 1991. Aging America. Department of Health and Human Services, Washington, DC.

DHHS, 2000. Healthy People 2010. Office of Disease Prevention and Health Promotion, US Dept of Health and Human Services, Washington, DC.

Diener, E., Emmons, R., Larsen, J., Griffin, S., 1985. The Satisfaction with Life Scale. Journal Personality Assessment 49, 71–75.

Duck, S., 1994. Meaningful Relationships. SAGE Publications, London.

Duck, S., 1995. Talking relationships into being. Journal of Social and Personal Relationships 12, 535–540.

Gelso, C., Hayes, J., 1998. The psychotherapy relationship: theory, research and practice. Wiley, New York.

Gilbertson, J., Dindia, K., Allen, M., 1998. Relational Continuity Constructional Units and the Maintenance of Relationships. Journal of Social and Personal Relationships 15, 774–790.

Grosz, B., Sidner, C., 1986. Attention, intentions, and the structure of discourse. Computational Linguistics 12, 175–204.

Hartley, A., 1992. Attention. In: Craik, F., Salthouse, T. (Eds.), The Handbook of Aging and Cognition. Erlbaum, Hillsdale, NJ.

Horn, J., 1982. The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in adulthood. In: Craik, F., Trehub, S. (Eds.), Aging and Cognitive Processes. Plenum Press, New York.

Horvath, A., Greenberg, L., 1989. Development and validation of the working alliance inventory. Journal of Counseling Psychology 36, 223–233.

Horvath, A., Symonds, D., 1991. Relation between working alliance and outcome in psychotherapy a meta-analysis. Journal of Counseling Psychology 38, 139–149.

Ice, G., 2002. Daily life in a nursing home: has it changed in 25 years? Journal of Aging Studies 16, 345–359.

Jarvis, K., Friedman, R., Heeren, T., Cullinane, P., 1997. Older women and physical activity: using the telephone to walk. Womens Health Issues 7, 24–29.

Lobach, D., Hasselblad, V., Wildemuth, B., 2003. Evaluation of a Tool to Categorize Patients by Reading Literacy and Computer Skill to Facilitate the Computer-Administered Patient Interview. *Proceedings of AMIA*, Washington, DC.

NHLBI, 1998. Clinical guidelines on the identification, evaluation and treatment of overweight and obesity in adults. Bethesda, MD: Department of Health and Human Services, National Institutes of Health, National Heart, Lung and Blood Institute.

NTIA, 1999. *Falling through the Net: defining the digital divide*. A report on thetelecommunications and information technology gap in America. Washington, DC: National Telecommunication and Information Administration.

Planalp, S., 1993. Friends' and acquaintances' conversations II: coded differences. Journal of Social and Personal Relationships 10, 339–354.

Planalp, S., Benson, A., 1992. Friends' and acquaintances' conversations I: perceived differences. Journal of Social and Personal Relationships 9, 483–506.

Pollack, M., Brown, L., Colbry, D., McCarthy, C.E., Orosz, C., Peintner, B., Ramakrishnan, S., Tsamardinos, I., 2003. Autominder: an intelligent cognitive orthotic system for people with memory impairment. Robotics and Autonomous Systems 44, 273–282.

Russell, D., Peplau, L., Curtrona, C., 1980. The revised UCLA lonelines scale: concurrent and discriminant validity evidence. Journal of Personality and Social Psychology 39, 472–480.

Smith, J., 2000. GrandChair: conversational collection of grandparent's stories. Massachusetts Institute of Technology PhD Dissertation, Cambridge, MA.

Tamura, T., Yonemitsu, S., Itoh, A., Oikawa, D., Kawakami, A., Higashi, Y., Fujimooto, T., Nakajima, K., 2004. Is an entertainment robot useful in the care of elderly people with severe dementia? The Journals of Gerontology 59A, 83–85.

Ulatowski, H., Cannito, M., Hayashi, M., Fleming, S., 1985. Language abilities in the elderly. In: Ulatowski, H. (Ed.), The Aging Brain: Communication in the Elderly. College-Hill Press, San Diego, CA, pp. 125–139.

Wada, K., Shibata, T., Saito, T., Tanie, K., 2003. Effects of Robot Assisted Activity to Elderly People who Stay at a Health Service Facility for the Aged, Proceedings of the 2003 IEEE/RSJ Intl Conference of Intelligent Robots and Systems, Las Vegas, NV.

Ware, J., Kosinski, M., Keller, S., 1996. A 12-item short form health survey: construction of scales and preliminary tests of reliability and validity. Med Care 34, 220–233.

Williams, M., Parker, R., Baker, D., Parikh, N., Pitkin, K., Coates, W., Nurss, J., 1995. Inadequate functional health literacy among patients at two public hospitals. JAMA 274, 1677–1720.

Woods, W., 1986. Transition network grammars for natural language analysis. In: Grosz, B.J., Jones, K.S., Webber, B.L. (Eds.), Readings in Natural Language Processing. Morgan Kaufmann Publishers, Inc, Los Altos, CA, pp. 71–88.