

Creating new technologies for companionable agents to support isolated older adults

CANDACE L. SIDNER, Worcester Polytechnic Institute

TIMOTHY BICKMORE, Northeastern University

BAHADOR NOORAIE, Worcester Polytechnic Institute

CHARLES RICH, Worcester Polytechnic Institute

LAZLO RING, Northeastern University

MAHNI SHAYGANFAR, Worcester Polytechnic Institute

LAURA VARDOLAKIS, Northeastern University

This paper reports on the development of capabilities for (on-screen) virtual agents and robots to support isolated older adults in their homes. A real-time architecture was developed to use a virtual agent or a robot interchangeably to interact via dialog and gesture with a human user. Users could interact with either agent on twelve different activities, some of which included on-screen games, and forms to complete. The paper reports on a pre-study that guided the choice of interaction activities. A month-long study with 44 adults between the ages of 55 and 91 assessed differences in the use of the robot and virtual agent.

Additional Key Words and Phrases: Content Indicators: I.2.m, I.2.9, H.1.2. keywords: virtual agents, virtual agent companions, social isolation, older adults, robot companions, situated dialog systems, real-time dialog systems, human-robot interaction, engagement, human-agent interaction studies, human-robot interaction studies.

ACM Reference format:

Candace L. Sidner, Timothy Bickmore, Bahador Nooraie, Charles Rich, Lazlo Ring, Mahni Shayganfar, and Laura Vardoulakis. 2018. Creating new technologies for companionable agents to support isolated older adults. *ACM Trans. Interact. Intell. Syst.* 1, 1, Article 1 (February 2018), 28 pages.

DOI: 0000001.0000001

1 INTRODUCTION

Social isolation is a broadly troubling trend in modern society. Many studies have demonstrated the negative effect of the lack of social support on health and well-being in the elderly; one study found social isolation leads to numerous health effects in elders (Nicholson 2012) while another found that mortality is significantly higher for elders who are socially isolated (Steptoe et al. 2012). Companionship and social support are also known to be significant positive factors in disease recovery and mortality, especially for older adults (Reblin and Uchino 2008) (Windle et al. 2011).

For purposes of this paper, social isolation is defined as the circumstances of living alone, having a small social network of friends and family, few social encounters and feelings of loneliness. Social support is defined as mechanisms that provide assistance to those who live in social isolation, either to reduce their actual isolation or to reduce their sense of loneliness or both.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2018 Copyright held by the owner/author(s). 2160-6455/2018/2-ART1 \$15.00

DOI: 0000001.0000001

What types of technology could support isolated older adults? Email, video calls and texting are already available, but social isolation continues as a troubling trend. Human companions are expensive, or when voluntary, are not available in the numbers needed. Intelligent virtual agents have been developed for many settings including home real estate, exercise promotion, touring cultural sites and children's education. An intelligent virtual agent serving as a social companion to an older adult (rather than say someone who makes a meal or takes a senior shopping) offers another technological possibility. To explore how adults serve as social companions, observations of adults who had volunteered to be social companions to older adults were conducted through a non-profit organization in Boston (Vardoulakis et al. 2012). Results from those observations indicated that volunteers saw their seniors as friends whom they visited 1-2 hours per week. Their activities including co-TV watching, small talk, senior story telling, sports, health of the senior, reports on the senior's family and senior future plans. To create a virtual agent that could serve as a friendly companion, the research team who undertook the work reported here considered ways in which an agent could perform companionable activities with a senior, although it was clear to the team that a virtual agent could not be counted as a friend in the human sense. At the same time, technology can be used to help seniors connect to other people. So the team also wanted to use the same virtual agent to assist seniors in keeping and making connections with others.

The focus of the research reported here has the potential for reducing isolation in older adults. When this research began, several questions guided the research team's efforts: What types of activities would interest older adults when presented on a computer screen, and what activities might reduce isolation? What technology could be used or developed to explore the use of both on-screen virtual agents and robots? How might that technology provide a familiar paradigm for interaction that was easy to use, presented the agents as sociable as possible, and could operate reliably for an extended period of time?

The project undertaken by the team is called the AlwaysOn project and was begun in 2011. It produced agents that are cognitively capable, meaning they have limited ways to reason about their human partners and their environment, and they can perform tasks behind the scenes to simplify some aspects of using computers that are frustratingly complex to many computer users. The agents are "Always-on agents," meaning that they are available to their human partner at any time the human user is present by means of various sensing mechanisms that are built into the agent. The agents use dialog as a means of communicating with their human partners, again to do away with much of the complexity of current computer (and cell phone) interactions. All these technologies are brought together in one system for the purpose of continuously supporting isolated older adults. AlwaysOn agents offer the potential to counteract isolation both directly, by providing interactions that approximate companionship, and as intermediaries, by putting isolated people in contact with other people, both electronically and physically. This paper reports on both the user studies and the technology built to use Always-on agents, both virtual and robotic.

2 RELATED EFFORTS

The research reported here grows out of work by Bickmore and colleagues, on developing virtual agent-based health coaching applications for older adults, primarily in exercise promotion (Bickmore et al. 2005, 2010, 2013). In these systems, on-screen agents were deployed to older adult homes for a month or longer, and the researchers found that older adult users could successfully use the applications and logged in regularly, resulting in significant increases in physical activity. The focus of the current work is on targeting social isolation, by providing a range of activities to involve the older adult, and in the technology infrastructure and morphology of the social agent interacting with the user.

A number of studies about the design of virtual agents for use with elders have been undertaken (Maciuszek et al. 2005; Tsourti et al. 2014; Yaghoubzadeh et al. 2013). Many of these efforts occurred well after the Always-on project had begun. These studies have considered types of tasks, whether a human-looking agent is valued, and constraints on what agents might do, including for elders with cognitive impairment. Among the tasks considered were ones to keep track of medications, play games, such as chess and provide reminders of appointments. These efforts were design studies only and do not report on attempts to build systems based on these studies. One system (Pinto et al. 2008) was designed to interact with older adults to do such tasks as newspaper reading and general chit-chat. However, no evaluation of this system with actual users was undertaken, nor was the system used in any significant way with older users. In contrast, a focus of the AlwaysOn effort was to develop a system that could be tested and could be used for more than a few days. A European Union project, *Miraculous Life*, (Miraculous Life 2013) posited the development of virtual agents to exist in the homes of elders to provide emotional support and connect the elder with caregivers and family on a daily basis. One report from the project outlines design considerations for such an agent (Hanke et al. 2015). The German-Greek project *Care* explored the use of digital display frames for recommender systems to provide older adults with wellness advice in their homes (Seiderer et al. 2015). They investigated activities that would support wellness for this group, built the technology and did prototype testing with two older adults who lived together.

The European Union VERVE project has also considered social isolation, especially for persons with Alzheimer's, Parkinson's disease, and phobia type anxieties, as well as older adults who fear falling. Their thrust has been to produce virtual worlds, and virtual environments in serious games with virtual agents to assist their target population. A challenging aspect of this research effort is to create very realistic characters in virtual environments. The VERVE team evaluated the use of serious games in patients with Alzheimer's and related diseases (Robert et al. 2014). In contrast, the efforts reported in the current paper are more modest because this work does not consider the challenges of aging adults with serious illness or phobias, and it does not undertake to create the rich types of worlds required in VR or serious games.

All the efforts cited above led our team to consider an agent who was able to communicate with a healthy older user, maintained the privacy of that user, and focused on computational capabilities that would support realistic communication and permit long-term evaluation of the resulting agent system with users.

Robotic technology has made substantial inroads in longitudinal care for older adults. The *Paro* seal-like robot has been used extensively in nursing homes to calm older adult patients (Wada et al. 2013), to provide pleasant pet-like interactions and to increase interactions among the older adults (Sabanovic et al. 2013). Researchers have used zoomorphic robots (Klamer et al. 2010) with a three-person sample of elders in their home for exercise coaching for a period of ten days; one subject used the robot faithfully but the others did not. Several European Union projects are focused on robots for elder care, especially mobile robots to help those with mild cognitive impairment or physical impairments (Johnson et al. 2014; Nani et al. 2005). The *Kompai* mobile robot, which started to be developed commercially in 2010, had an architecture that included providing games, weather talk, health checks, a shopping list and an agenda for elderly users (Guizzo 2017). Testing of *Kompai* with speech input concluded that both vocabulary and syntax needed to be greatly increased while the testing of the computer screen icons yielded mixed results; no long-term studies of the robot with elders is available (Grenata et al. 2010). A comparison study of *Kompai* with another mobile robot (operated in *Wizard of Oz* mode) with elders (Huijnen et al. 2010) indicated that users expectations for the robots exceeded their capabilities, that companionship acceptance depended on knowing when to react and smart dialog management, while privacy and technology

acceptance depended on transparency of the technology, social situation awareness of the robot and helpful responses. Other efforts for robots and the elderly include teleoperated robots to help with everyday tasks (Mast et al. 2012). The AlwaysOn project learned from these efforts about the challenges of speech, the need for good dialog, and that privacy matters would need to be considered in developing a system. The AlwaysOn project also chose to include walking promotion as an activity for users.

A recent study (deGraaf et al. 2016) reports on the attitudes and use of the same zoomorphic robot as (Klamer et al. 2010) but greatly enhanced to broadcast news, messages, music, texts, alerts, and radio and with a built-in webcam to enable users to communicate with family at home or to watch their homes when absent. This robot was placed in 70 homes with a total of 102 adults (a mixture of single adults, families with young children and families with teenagers) for 6 months. The researchers considered usage during six phases: pre-adoption, adoption, adaptation, incorporation and identification, which the team devised based on other research. Of the total participants, 55 participants used the robot through all the phases of usage, while 21 stopped at the incorporation phase of use. For adaption through indentification phases, the researchers found that usage tended to drop except for a group of dedicated participants in the study. Participants were asked about a large number of attitudinal and social beliefs including usefulness, enjoyment, adaptivity, trust and intelligence (which dropped over the 6 phases), while sociability, attitude towards robots, and companionship rose. The AlwaysOn project is another effort to understand the value of robots and virtual agents in extended home use, albeit with a different population than the zoomorphic robot project.

Studies have compared GUI interfaces, robots and on-screen agents. (Fasola and Mataric 2013) created a social robot, that is, a humanoid robot with speech ability and which performed social behaviors including praise for effort and humor/commentary, for exercise coaching with older adults. In their evaluations using either the physical robot or a virtual (on-screen) version of the robot, subjects rated the physical robot more highly on measures of enjoyability, social presence, usefulness and social attraction. (Kidd and Breazeal 2004) concluded that robots were favored over onscreen characters, while (Wainer et al. 2013) concluded that users preferred a physically present robot (their robot had a head, but no face or other humanoid features besides movement) to the same one presented on over a video link versus a simulation. However, (Wrobel et al. 2013) found that for elders performing a game task, the preference was for a standard computer screen followed by a robot, and lastly by an on-screen agent. That work also pointed out that the elders indicated that they were focused on task performance, and so having fewer new devices were better for getting work done. Hasagawa and colleagues (Hasegawa et al. 2010) found when comparing a physical robot with humanoid features, an animated on-screen agent character of the same robot, and a GPS system for the task of getting directions, users dispreferred the GPS system over the robot and screen agent. However, users performed more gestures with the on-screen agent than with the robot. These various efforts encouraged our team to include a robot as an AlwaysOn agent, but to otherwise hold fixed how both agents interacted with users.

3 INITIAL USER STUDIES WITH ELDERLY

To better understand what users wanted to do with an agent, the effect of being always on, and to evaluate some candidate activities, several of the authors conducted a series of preliminary studies. These studies included a Wizard of Oz (WOZ) study in which several users interacted with a virtual agent controlled remotely by an experimenter, and an early prototype of a fully autonomous virtual agent. Both studies involved having the agent in the homes of older adults for a week.

The purpose of the WOZ study (Vardoulakis et al. 2012) was to determine the range of topics that isolated older adults would want to discuss with a virtual agent in their home. Participants were 12 older adults (10 women, 2 men) who lived alone, aged 56-73, generally well-educated (all but one had some college). They scored an average of 38.6 (SD 8) out of a possible 80, on the UCLA loneliness instrument (Russell and Peplau 1980), indicating overall low levels of loneliness. More than half of the participants wanted to talk about the weather, family and personal stories (telling stories to the agent). Additionally, half of all the conversations involved discussing the participant's future plans and activities.

The purpose of the second study was to determine the effectiveness of proactive conversation initiation by a virtual agent. This study involved a fully autonomous virtual agent (Ring et al. 2015) that could talk to elders about a limited range of topics, including chat about the weather and exercise, and humorous anecdotes told by the agent. This system used a motion sensor to detect the user's presence and initiate conversations when the elder walked near the agent. Twelve (1 Male, 11 Female) older adults, aged 56-75, who lived alone, participated in the week-long study, and were randomized either to the sensor-driven proactive system or an equivalent system that did not proactively initiate conversations. Participants conducted 15.9 (SD 8.1) interactions per week on average, lasting an average of 140 (SD 26) seconds each. Participants in the proactive condition reported feeling significantly less lonely, $t(154) = -3$, $p < .05$, and more satisfied $t(154) = -4.04$, $p < .05$, compared to participants in the non-proactive condition.

4 DESIGNING FOR THE USER EXPERIENCE

The outcomes of the previously discussed studies influenced the design of the user experience with the agents of the AlwaysOn system in terms of types of activities provided to the planned population of older adults and in terms of the paradigm of conversational interaction, especially dialog. The type of agent, which is discussed below, was one the research team wished to study for its effects on user behavior and satisfaction.

The main interaction paradigm in the AlwaysOn system is conversation, and in particular, dialog. The agent makes its contributions to the dialog using speech and gestures, and the user chooses his/her contribution from a menu of utterances provided on the touch screen. Dialog is thus the principal way of interacting. Dialogs occurred around various activities, and users could undertake each activity for an extended period of time if desired, from a few minutes to one-half hour. Some of the activities involve additional on-screen graphics, such as the card game shown in Figure 1, where the user directly manipulated the cards on-screen. However, the team otherwise eschewed other traditional GUI methods using icons, pull-down lists, clicking on apps, and the like, in favor of using speech output and menu input dialog interaction, as illustrated in that figure. This choice provided an interface with relatively few on-screen demands in order to interact with the agent and to perform activities.

The agent speaking and the user responding via text offered an interaction that was not quite a normal face-to-face conversation, but one very close to it. The exceptions to this conversational style were direct manipulation, which was needed for playing cards or tic-tac-toe, and a virtual keyboard to allow typing in new proper names of people and places. The research team's motivation for this design choice was to provide a single interaction method that would seem familiar to users and would require the least additional mechanisms. The team chose not to engage in speech understanding by the agents because 1) voice models for older adults are still not widely available in speech recognition; 2) Bickmore and colleagues (Bickmore et al. 2013) had used the same type of interface with 132 geriatric patients for two months for exercise coaching, and none of the participants had any trouble using the interface; 3) other research (discussed earlier) suggested



Fig. 1. The on-screen agent Karen plays cards with the user.

speech was a technical challenge; and 4) progress in speech understanding was not a focus of the project.

Two forms of embodiment of a social agent were developed. Karen, shown in Figure 1, comes from the work of Bickmore et al. 2005. Karen is a human-like agent animated from a cartoon-shaded 3D model. She is shown in Figure 1 playing a social game of cards with user. Notice that user input is via a touch-screen menu using an ASUS touch screen computer. Also, the speech bubble does not appear in the actual interface, which uses text-to-speech generation. The Reeti robot (Reeti 2017) shown in Figure 2 is a desktop robot with a height of approximately 2 feet. The robot has cameras in its eyes, a moveable mouth, 2 degrees of freedom in its neck, and “ears” that can rotate. The robot used the ASUS touch screen for all the user responses to the robot, exactly as the user did for the virtual agent. In user interactions both agents were referred to as Karen.

Sensor technology for the agents is identical. Both agents, in addition to speech generation using Ivona, Salli voice, produce nodding gestures in the dialog as a form of backchanneling (Yngve 1970). They also track the user’s face using off-the-shelf face recognition software, either through an onboard camera in the ASUS computer, for the virtual agent, or through the robot’s eye-based cameras.¹ They both used a very simple infra-red motion detector for sensing motion in the room. The motion sensor is the small black box, outlined in red, with the white button visible in the lower left corner of the ASUS interface screen in Figure 2. When an activity provided information on the computer screen, the agent generally turned to look at the screen after its contribution to the dialog. Five of the activities provided some on-screen information that the user manipulated.

¹A video example of the robot interacting with a person on a number of activities is available at: (Sidner et al. 2015).

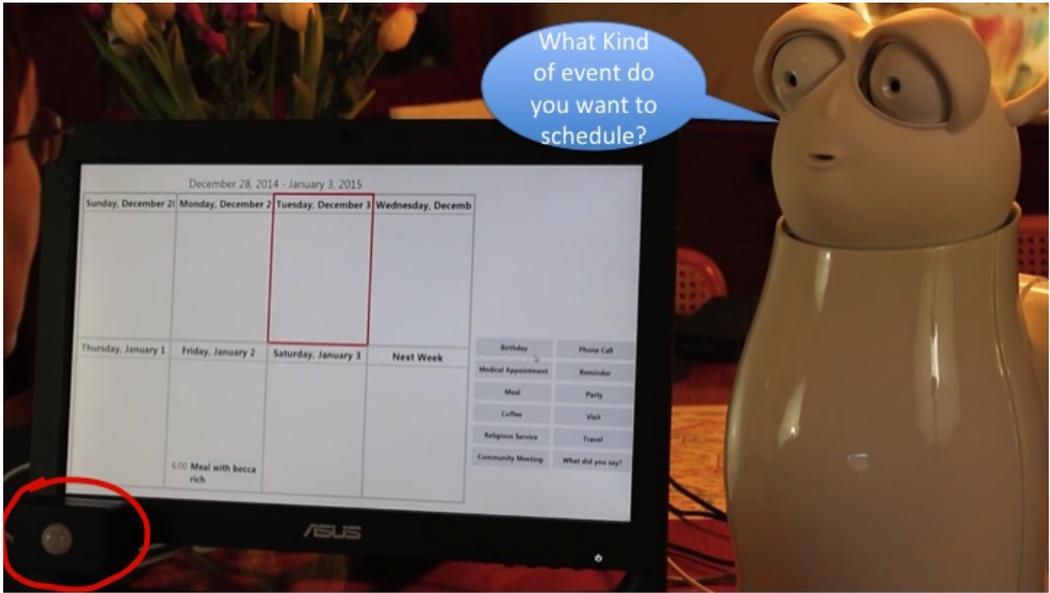


Fig. 2. The robot Reeti interacts with the user and the calendar activity.

The type of robotic agent used in AlwaysOn resulted from practical concerns. While the team would have preferred a more human-looking robot (to match and compare more readily with the human-like on-screen agent), reliable human-looking robots that could operate for the planned time frame of one month of use without failure were not available. As the team was planning on making the AlwaysOn system available for a full month with older users, Reeti was the only commercially available robot that had been used by other groups reliably for extended periods of time.

The project uses both an on-screen agent and a robot in order to understand what differences might exist in their use over a long period of time. A number of studies have shown that users trust robots more than on-screen agents (c.f. the studies mentioned above), but most of those studies did not attempt to involve users for more than a single interaction. A notable exception is the work of (Kidd and Breazeal 2008) on long-term interaction with a wait-loss coach robot. However, they did not compare on-screen agents and robots, just robots to computer logging. The research team imagined that in long-term use, users might not find one more preferable than the other, even though certain behaviors, such as face tracking are more noticeable with the robot.

The two agents were programmed to offer and participate in an array of activities with the user. Different activities can serve different goals, either improving the user’s well-being by doing some socially engaging activity or connecting users to friends and family beyond the agent interface. A reasonably extensive number of activities were needed if users were to interact over many days

and weeks without getting bored with the system. The initial studies undertaken by the authors indicated that users wanted more things to do with the agent when the interactions took place over several days. This result pushed the AlwaysOn team to devise a number of different types of activities.

Activities such as **talking about the weather** or **playing a social game of cards** are meant as ice breakers, to give the user to get to know the agent. Social game play was created by developing an game architecture and sub-system that would allow the agent to talk about how the game was going (Behrooz et al. 2014). Three games were developed using this architecture: the card game rummy, classical checkers and tic-tac-toe. Each of these games provided a direct-manipulation interface to make it possible to play the game. Activities to support well-being included getting **health** tips and **nutrition** coaching, which also presented charts about the user's progress and forms to fill in.

Other activities were more instrumental in nature, such as the agent's **about** self-introduction dialogs, in which it interactively explained its capabilities to the user. For each user, an initial **enrollment** dialog, in which the person's name, birthdate and friends and family names were requested, gave the agent some initial information about the user. Once the agent knew about the user's family and friends, one of its activities supported the reported user desire from our initial studies to **explain** about family and friends, albeit in a very limited fashion. Other instrumental activities included a personal **calendar** activity (shown in Figure 2 with the Reeti robot), which presented a calendar with events and was used to add or delete events such as birthdays and plans with friends. The agent made changes based on dialog with the user. To provide some light-hearted enjoyment, the agent told short humorous **anecdotes** to the user, an activity that users reported liking in preliminary studies. All these activities made for a social agent who was useful and yet could make light-hearted conversation with the user.

Several activities directly address the project goal of socially connecting the user with other people. The team implemented *Video buddy*, an activity in which the agent arranged **video calls** for the user with family and friends. In this activity the agent undertook all of the complexity in standard video interfaces for contacting the video partner and setting up the call. The user simply indicated that he or she wanted to reach a person and the rest was undertaken by the agent. The user did not have to click on various icons presented on the left and right of the screen to determine how to make a call (as is the case with SKYPE and Hangout), nor did the user have to deal with calls that failed in some way. The agent was programmed to indicate to the user that a call ended prematurely if video was lost. While the agent set up the call, the user was free to do other activities, and the agent would interrupt the user when the call was ready to take. The activity to coach outdoor **exercise**, an area well explored by Bickmore et al. in earlier work (Bickmore et al. 2005), had been shown to result in more social connection by getting the user out of his/her dwelling and into the community (Bickmore et al. 2013). A sample interaction of exercise coaching is shown in Figure 3.

Finally, the personal life **story** acquisition activity addressed both a project goal and a benefit for seniors. Seniors tell stories about their lives in part to make sense of their full life experience (Butler 1964). In the life story acquisition activity, the user told a personal story, which was audio and video recorded by the agent for later sharing with the user's family and friends over the Internet; this activity offered mitigation of social isolation by providing the elder with a new means of connecting with others. Life story acquisition was supplemented by a series of suggestions of positive topics, one of which the agent could suggest to the user in case the user could not him or herself decide how to organize his or her storytelling thoughts.

User chooses from activities menu: Discuss Exercise
Agent: So let us talk about how you did on your walking. So the last time we talked you said you would go for a walk. Did you get a chance to go for a walk yesterday?
User: Yes, I took a 20 minute walk.
Agent: Do you think you might do anymore walking today?
User: Yes, I might do some more walking today.
Agent: Looks like you got your walking in. Wonderful. We are doing some great work together. So can you go for a 5 minute walk before we talk again?
User: I can go for more than that.
Agent: How many minutes can you walk?
User: 20 minutes.
Agent: Are you sure you can do that much? I would suggest taking it slower at first.
User: I am sure I can do that.
Agent: That is awesome. We make a great team. So let us review how you are doing.
User: OK.
Agent: Here is a snapshot of how well you have been doing. Here it is.
<display of chart with minutes walked versus number of days>
As you can see on the chart, a perfect job on your exercise.
User: OK. I am finished looking at this.
Agent: I know you will do great. Ok, now, what would you like to do together

Fig. 3. A conversation with the agent about exercise

A typical interaction with the agent might start with some greetings (specific to the time of day) and then some discussion of the weather. The weather discussion could be as short as today's weather forecast or extend to the next day, weather in other cities, and weather where friends or family live. At the user's choice, weather might be followed by a social game of cards where the agent's and user's hands in the game and the way the game is played out are commented upon. Additional activities would follow based on what the user had learned about the set of activities available and how long the user had interacted with the agent. As the user had more sessions, additional activities were introduced thereby reducing overwhelming the user with all the activities early in the user's experience the with agent.

A sample conversation with the agent when first interacting with an agent for the day, but when the time of day is past noon, is shown in Figure 4.

The AlwaysOn system continuously maintained a model of the state of engagement (Sidner et al. 2005) between the user and the agent. The agent remained quiescent, with no gestures except eye blinks for the virtual agent and eye movements for the robot agent, until it sensed nearby motion (via an infrared motion detector) followed by the appearance of a face in its vision system. Given that information, the agent would look at the user and offer a greeting. No response returned the agent to the quiescent state. A user response to an agent greeting moved the agent to the engaged state. Engagement was assumed successful as long as the user remained in view of the agent and responded in the dialog. The disengagement state could come about at the natural conclusion of the conversation (with goodbyes), or when the user was no longer visible for an unexpected reason for an extended period of time, e.g., to answer a ringing doorbell. Because the AlwaysOn agents could not understand sounds in the environment, they could not know why the user might be disengaging, but they had simple strategies for dealing with what they perceived as unexpected disengagement. They waited for a period of time after asking "are you still there?" If the user did

User: Hi.
Agent: How's your day going, Joe?
User: Fine.
Agent: Good glad to hear it. Ok, now how's the weather?
User: I'd like a weather forecast.
Agent: Today it is going to be 75.6 Fahrenheit and sunny.
User: What is the forecast for tomorrow?
Agent: Tomorrow it will be sunny, low around 60 F. Winds 6 to 10 miles per hour.
User: Stop discussing the weather.
Agent: Ok, now what would you like to do together?
User: <list of activities>

Fig. 4. A sample beginning of the day conversation

not answer and did not return, the agent would disengage and end the interaction. The agent did not initiate disengagement, although it could attempt to hurry the conclusion of a session if some event in the user's calendar was about to start.

5 AN ARCHITECTURE TO SUPPORT AGENTS WITH ONGOING ENGAGEMENT AND DIALOG

The architecture of the AlwaysOn system was driven by the need for ongoing engagement and dialog with its human user. The research team started out using the collaborative dialog system Disco system for dialog (Rich and Sidner 2012), but it became evident that the proper integration of dialog and gestures, which supported dialog and engagement, required more real-time control. An additional software challenge was allowing the activities about which the user and agent had dialogs to be built by many members of the team and with different software tools, but to be easy to integrate into the AlwaysOn system. Thus modularity and extensibility with respect to activities was a necessity (Nooraie et al. 2014). This section discusses both of these concerns with specific needs that are enumerated below.

5.1 Use Cases for Real-time Engagement and Dialog Interaction

A detailed set of **use** cases, a methodology from software engineering, guided the development of the AlwaysOn system for its engagement and dialog needs. In this approach, one identifies at the beginning of the design process a collection of archetypal behaviors (the "use cases") that the system should support and then evaluates the implemented system in terms of how well it supports each of these behaviors. Furthermore, the use cases should be chosen to cover the most challenging aspects of the system's required behavior. The nine cases presented here support engagement and dialog between the agent and the user. As defined by Sidner et al (Sidner et al. 2005), engagement is "the process by which two (or more) participants establish, maintain and end their perceived connection during interactions they jointly undertake." Furthermore, as shown below, engagement involves both verbal and nonverbal behavior.

Case 1) Walking Up to the Agent: The first use case relates to establishing engagement. Specifically, when the user walks by and triggers the motion detector, the agent should awake out of its quiescent state and issue a greeting, such as "Good morning." If the user then responds by approaching the computer (which the agent can notice with face detection software on the webcam), the agent should continue start face tracking (see next use case) and continue with a further engaging utterance, such as "How did you sleep?"

The next four use cases relate to maintaining engagement. Related work identified four types of “connection events” that function to maintain engagement (Holroyd et al. 2011; Rich et al. 2010). Each type of connection event has a corresponding use case. In general, these use cases involve crucial timing constraints between the verbal and/or nonverbal behaviors of the user and the agent.

Case 2) Face Tracking: Face tracking is the agent’s attempt to achieve what is technically called mutual facial gaze (Argyle and Cook 1976) with the user. When the agent is face tracking, it should orient its gaze toward where it detects the user’s face. In addition to being the agent’s default gaze behavior for maintaining engagement, mutual facial gaze can have other interaction functions. For example, it is typical to establish mutual facial gaze at the end of a speaking turn (see next use case).

Case 3) Turn-Taking: Even though a conversational interaction entails much more than turn-taking, an embodied conversational agent nevertheless does need to manage speaking turns, particularly in a menu-based system. In linguistics, an adjacency pair (Schegloff and Sacks 1973) is the term used to refer to two utterances by two speakers, with minimal overlap or gap between them, such that the first utterance provokes the second utterance. A question-answer pair is a classic example of an adjacency pair. Thus, after producing the so-called “first turn” of an adjacency pair, the agent should wait until the user responds (or until some specified timeout occurs). In some conversational circumstances, the user’s response can also be followed by a “third turn” in which the agent, for example, acknowledges the user’s response.

Importantly, the concept of adjacency pair is generalized beyond the traditional linguistic definition to include both verbal and nonverbal responses. So for example, a nod can be the answer to a question, instead of a spoken “yes,” or the performance of an action can be the nonverbal response to a verbal request, such as, “please pass the salt.” Adjacency pairs, of course, also often overlap with the other nonverbal behaviors, such as face tracking and directed gaze (see Use Case 5).

Case 4) Backchanneling: A backchannel is an interaction event in which a listener directs a brief verbal or nonverbal communication back to the speaker during the speaker’s utterance. Typical examples of backchannels are nods and/or saying “uh, huh.” Backchannels are typically used to communicate the listener’s comprehension of the speaker’s communication (or lack thereof, e.g., a quizzical facial expression) and/or desire for the speaker to continue. A conversational agent should be able to both generate appropriate backchannels and interpret backchannels from the user.

Case 5) Directed Gaze: Finally, Figure 5 shows the time line for the last, and most complex, engagement maintenance use case, called directed gaze (Kendon 1967). In this behavior, one person (the initiator) looks and optionally points at some object in the immediate environment in order to make it more salient, following which the other person (the responder) looks at the same object. This behavior is often synchronized with the initiator referring to the object(s) verbally, as in “now spread the cream cheese on the cracker” (pointing first to the cream cheese and then to the cracker). By turning his gaze where directed, the responder intends to be cooperative and thereby signals his desire to continue the interaction (maintain engagement).

In more detail (see Figure 5), notice first that the act of pointing (1), if it is present, begins after the initiator starts to look (2) at the object. Looking at the object is likely because it is hard to accurately point at something without looking to see where it is located. After some delay, the responder looks at the salient object (4). The initiator usually maintains the pointing (1), if it is present, at least until the responder starts looking at the object. However, the initiator may stop looking at the object (2) before the responder starts looking (4), especially when there is pointing. This switch can occur so that the initiator can check whether the responder has directed his gaze yet. Finally, there may be a period of shared gaze, i.e., a period when both the initiator (3) and responder (4) are looking at the same object.

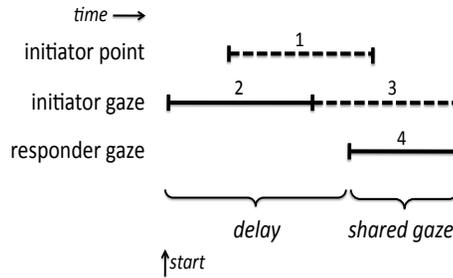


Fig. 5. Time line for directed gaze

Case 6) *Walking Away*: This use case relates to ending engagement. Hopefully, most of the time disengagement between the agent and user will occur as the result of an explicit leave-taking conversational exchange, such as, “Goodbye; see you later.” However, the agent should also be prepared to deal with the user simply walking away at any time.

The remaining three use cases relate to various kinds of interruption behaviors that are essential to dialog behavior. The ability to do more than one thing at a time and smoothly shift between them is a key requirement for a natural conversational agent.

Case 7) *Scheduled Event*: One reason for interrupting an activity is to due the (imminent) occurrence of a scheduled external event. For example, in the AlwaysOn effort, the agent helps the user keep a personal calendar of events such as lunch dates, doctor appointments, etc. If the user has a lunch date at noon, the agent will interrupt whatever the agent and user are doing together (e.g., playing cards) ten or fifteen minutes before noon to remind the user of the lunch date and to wrap up or postpone the current activity.

Case 8) *Changing Topic*: A conversational agent should be able to, either of its own volition, or in response to the user’s behavior, smoothly change the topic of the conversation and then, if appropriate, smoothly return to the original interrupted topic. For example, in the AlwaysOn effort, activities such as playing cards are viewed as social “containers,” within which other topics can also be discussed. For example, at the end of the user’s turn in a card game, the agent could say, “By the way, have you thought about my suggestions for how to get more exercise?” After a short discussion of exercise, the agent could return to the card game by saying, “Ok, so I think it’s your turn now.”

Case 9) *Barge-In*: Barge-in is a common dialog phenomenon similar to backchanneling (Case 4), in that the listener starts communicating before the speaker’s turn is finished. In the case of barge-in, however, the listener’s intention is for the speaker to stop and let the listener “take” the turn. A conversational agent should be able to respond to the user’s barge-in by promptly ceasing to talk. In a purely spoken language system, the user can barge in simply by starting to speak. A menu-driven system, such as the one used in the AlwaysOn project, can support user barge-in by displaying a menu for the user to click on while the agent is still speaking. Should the user click on the menu item, the agent should stop speaking its turn at that time. The case of a conversational agent barging in on the user is less common. However, Chao et al. (Chao et al. 2011) describe a strategy, called minimum necessary information, for a robot to start acting before the user has finished speaking instructions.

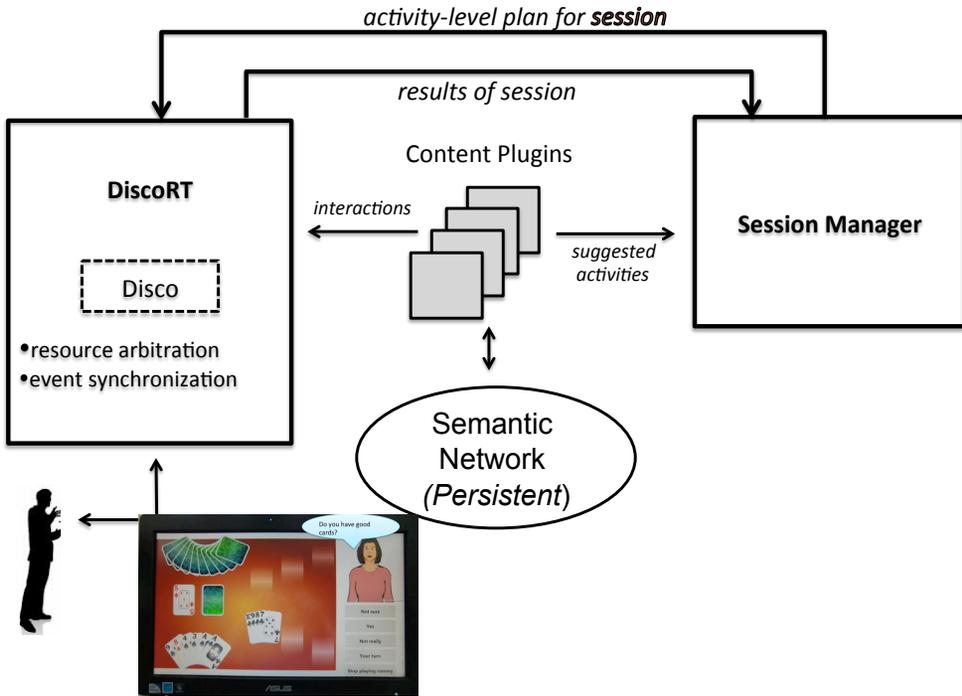


Fig. 6. The overall architecture of the AlwaysOn system

Figure 6 shows the overall architecture of the AlwaysOn system in which both hard real-time (taken to be responses in milliseconds) and long-term time (days or weeks) could both be supported. The solution was an extension, called DiscoRT (Disco for Real-Time) on the left side of the figure to focus on real-time needs. For the long-term needs, the Session Manager, shown on the right side of 6, organized which activities would be undertaken at each session with the user as an activity plan. Individual activities were authored as plug-ins to the architecture. Plugins thus allowed for modularity and extensibility of activities. Plugins were made available to the Session Manager to use to determine a plan as a combination of all the activities and of the results of previous sessions, during which the user could try a new activity or update a previous activity with new information or events. It also used general rules on how to introduce the user to yet more activities until all the activities were available to the user. While the option of allowing users to indicate activities they never wanted to do again or do at every session was not explored, the Session Manager would be the locus of such choices.

In brief, DiscoRT implements an arbitration-based parallel schema architecture that handles such hard real-time phenomena as barge-in (the agent immediately stops speaking in the middle of an utterance when the user touches a menu item on the screen) and time-outs (the agent repeats an utterance when the user did not respond after some time). DiscoRT also coordinates the inputs from the agent’s sensors and helps manage Disco’s dialog focus. Hard real-time was a critical feature of the AlwaysOn system not only for barge-in and time-outs, but also to keep the agents tracking

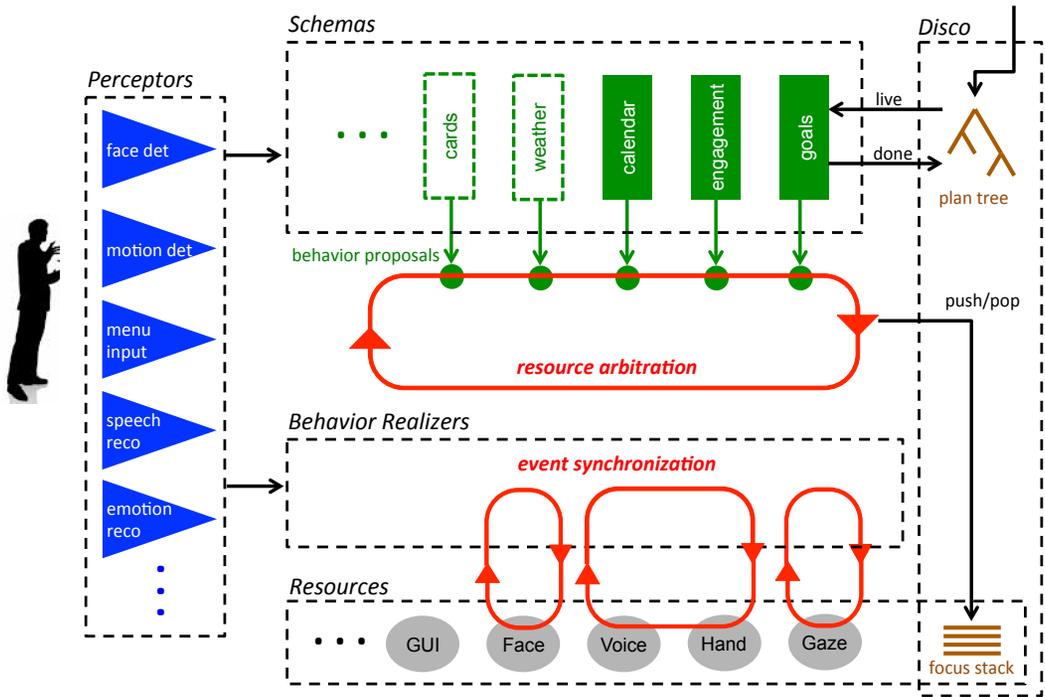


Fig. 7. The detailed architecture of DiscoRT

the user (via eye gaze) as the user moved about, to maintain engagement, to participate in dialog, and to respond quickly in game playing, calendar use and other interactions that required direct manipulation of the screen.

The design of DiscoRT also aimed to support virtual agents and robots that, unlike in the AlwaysOn system, can use their hands and arms to point at and manipulate objects in their environment. DiscoRT is designed to support fully spoken interaction, as well as the menu-based interaction mode of the AlwaysOn effort.

The Semantic Network in Figure 6 stores information gleaned from the user about his/her interests and activities, as well as basic models of such events as time for use in calendaring and the like. The AlwaysOn system is implemented in a combination of Java and .NET on Windows.

Figure 7 shows the detailed, abstract machine architecture of the DiscoRT system that was designed, implemented and used for the AlwaysOn agents to support the use cases above in a principled and general way. A discussion of the details of the architecture will provide the basis for accounting for how the nine use cases have been implemented. Four key features of this architecture are:

- 1) Multiple Threads: Supporting the continuous mutual signaling model of interaction obviously requires a highly parallel architecture with multiple threads. Each input and output modality and the internal decision-making processes needs to function independently without blocking one other.
- 2) Resource Arbitration: The agent’s face, voice, hands and gaze are resources that can be used for different (and sometimes competing) purposes in an interaction. For example, the agent’s gaze can be used to achieve mutual facial gaze (Case 2) or it can be used direct the user’s gaze to a

salient object in the environment (Case 5). Similar to a computer operating system, one of the key functions of DiscoRT is to arbitrate between competing demands for a given resource.

3) Real-Time Control: Timing is critical in all of the use cases. However, there is more margin for error in some use cases than in others. For example, an inappropriate delay of even a fraction of a second in the agent's response to a user-initiated directed gaze or barge-in would be noticeable and degrade the believability of the agent. On the other hand, changing topics or reminding the user of an upcoming scheduled event can be delayed a second or two without harmful effects. These different timing needs are dubbed "hard" and "soft" real-time, respectively, and handled separately in the architecture.

4) Dialogue Management: A unique feature of DiscoRT architecture is its integration with the *Disco* dialogue manager, shown in Figure 7 on the far right. Among other things, *Disco* provides a focus stack for helping to manage interruptions.

The left of the figure notes **perceptors** which is the system's abstraction for sensing capabilities, such as face detection, motion detection, menu input, speech recognition, emotion recognition and so on. Some perceptors simply abstract the output of a single hardware sensor, such as the infrared motion detector in the *AlwaysOn* project. However, perceptors can also fuse information from multiple input modes, such as an emotion recognition perceptor that combines facial expression information from a camera with tone of voice information from a microphone. Both the schemas and the realizers use the output of the perceptors. For example, the schema that establishes engagement uses face detection and motion perception. Note that there is no emotion recognition used by *AlwaysOn* agents, but the architecture is general enough to include such behavior.

Schemas are the core of the DiscoRT architecture. For the *AlwaysOn* agents, there is a schema corresponding to each social activity that the agent can do with the user, such as talking about the weather, playing cards, etc. These schemas are created and destroyed as the system runs. A few other schemas, such as the schema that manages engagement and the one for goals, are always running. The fundamental function of a schema is to continually propose behaviors. Each schema runs on its own thread with a loop cycle rate depending on its needs. To decide what behavior to propose at a given moment, a schema can consult its own state variables and the system's perceptors, as well as external sources of information. For example, the weather schema downloads up-to-date weather information from the Internet. Schemas make behavior proposals, which are programs to be executed. For example, the directed gaze behavior (see Figure 5) requires three resources (voice, hand and gaze) and includes a complex realizer program to synchronize the control of the resources with input from the perceptors. There are no restrictions in DiscoRT on the internal implementation of a schema, and the *AlwaysOn* schemas vary from state machines implemented in Java to programs written in the specialized language for *Disco* programming.

The **resource arbitration** thread/loop runs approximately once per second and gathers up the collection of behavior proposals from all of the running schemas. Proposals with non-overlapping resource requirements are directly scheduled, i.e., an instance of the specified realizer is created and started running (unless it is already running). Resource conflicts between proposals are resolved using a simple system of per-schema priorities with some fuzzy logic rules (Fodor and Roubens 1994) to make the system more stable, i.e., to prevent switching behaviors too quickly between closely competing schemas. The behavior proposals that are chosen are then scheduled.

Since DiscoRT is designed to support conversational agents, it includes specialized machinery for **dialogue management**. It uses the *Disco* dialogue manager, an open-source successor to Collagen (Rich and Sidner 1999) (Rich et al. 2001). *Disco* has two key data structures: a plan tree and a focus stack. Each of these has a point of integration with DiscoRT, as shown in Figure 7.

The plan tree, which is typically provided from another component outside of DiscoRT, (in the case of AlwaysOn the plan tree is the activity plan produced by the Session Manager in Figure 6) represents the agent's goals for its interaction with user as a hierarchical task network (Rich 2009). This formalism includes optional and repeated goals and partial ordering constraints between subgoals. Thus the typical interaction example, starting with greetings, etc., described in Section III above is formalized as a plan tree. The plan tree can be updated while the system is running.

DiscoRT includes a predefined schema, called the goals schema in Figure 7, which is always running and automatically starts the schema(s) corresponding to the currently live goal(s) in the plan tree. Thus, for example, when the "discuss the weather" goal becomes live, the goals schema starts the weather schema. When the schema exits, the corresponding goal in the plan tree is automatically marked as done. The focus stack is stack of goals, which captures the familiar phenomenon of pushing and popping topics in human conversation. For example, the agent can interrupt the user's activity to inform him or her of a scheduled meeting, but if the user wants to continue the current activity, the focus stack makes the return to the activity possible.

In DiscoRT, interaction that facilitates changing goals is achieved by making the focus stack a resource that represents control of the current topic of conversation. Schemas that involve conversation, such as the cards schema, the exercise schema, and the calendar schema in the example above, require the focus stack in their behavior proposals. When the resource arbitrator starts a new behavior realizer that requires the focus stack, it pushes the goal associated with the proposing schema onto the stack (unless it is already there). When the behavior realizer finishes, the goal is automatically popped off the stack, as is discussed for use case 7. This integration between the dialogue model and the schema architecture in DiscoRT is powerful and flexible. For example, it is possible and sometimes useful for a schema to propose a speech behavior without requiring the focus stack. If you poke a robot, it might respond by saying "Ouch!" without changing the conversational focus. DiscoRT also was provided with hooks for automatically producing generic transition language when the stack is pushed and popped, such as "excuse the interruption, but. . ." or "now, returning to..."

The **behavior realizers** (bottom of Figure 7) realize the gestures and verbal responses of the agent. Realizers implement the hard real-time event synchronization in the system; hard real-time events are those that require nearly spontaneous responses to user input, such as barge-in, time-out and responding to user-initiated directed gaze. A behavior realizer in DiscoRT is very similar (hence the name) to a BML realizer (Vilhjálmsson et al. 2011). However, for the reasons discussed in detail in (Holroyd and Rich 201), this architecture uses an event-driven Petri net rather than a fixed schedule as in most BML realizers.

Further, in DiscoRT there can be multiple realizers independently controlling non-overlapping resource sets (think of a robot "rubbing its tummy and patting its head" at the same time). Each behavior realizer has a separate thread that runs by default at 10Hz. Realizers often get information from perceptors, The DiscoRT ones support all of the essential timing relationships of BML (synchronize, before, after and fixed delay). The complete BML specification was not implemented because much of it concerned specific gestural actions; DiscoRT uses a more general framework. Also, the realizer programming API for DiscoRT uses Java rather than XML.

5.2 Implementation of Use Cases

With the details of the architecture in mind, the use cases as applied to the AlwaysOn architecture and agents can be discussed in more detail.

Case 1) Walking Up to the Agent: The engagement schema (implemented as a state machine) continually polls the motion perceptor (which abstract the infra-red motion detector hardware).

When motion is detected, the schema proposes a speech behavior (e.g., “Good morning” before noon and “How’s your day going?” between noon and 6 p.m.) and enters a state in which it starts polling the face perceptor (which abstracts the output of face detection software operating on the webcam output). When a face is detected, the schema proposes another speech behavior and enters the engaged state; otherwise, after a timeout, the schema returns to the motion perceptor polling state.

Case 2) Face Tracking: A behavior realizer that requires the gaze resource and uses the face detection perceptor implements face tracking. The realizer simply runs a loop that updates the agent’s gaze to where it currently sees the user’s face. The face tracking behavior that causes the realizer to be started is proposed by the engagement schema when it enters the engaged state (see Case 6 for stopping the realizer).

Case 3) Turn-Taking: Turn-taking is implemented by an abstract state machine that is reused in all of the schemas that include conversational behavior, such as weather, cards and calendar. The agent’s utterances are produced by proposing speech behaviors. In menu-based systems, the state machine waits for the user’s response by waiting for an event from the menu perceptor (which abstracts the menu GUI).

Case 4) Backchanneling: Backchanneling is a hard real-time phenomenon, so it must be implemented by a behavior realizer that receives events from a perceptor that detects appropriate moments at which to produce backchannels, such as the end of phrases or sentences in the user’s speech. The schema that proposes the backchanneling behavior is responsible for deciding what form of backchannel should be used (e.g., positive or negative) and for re-proposing a new behavior when the form of backchannel should be changed.

Case 5) Directed Gaze: As discussed previously, the time line for directed gaze (Figure 5) is implemented as a behavior realizer. Any schema can propose a directed gaze behavior.

Case 6) Walking Away: The engagement schema state machine includes a timeout to notice when there has been no face or motion perceived. When this case occurs, it stops proposing face tracking, asks if the user is still available and if not, enters the waiting for engagement state described in Case 1.

Case 7) Scheduled Event: While the user and agent are playing cards together, both the card schema and the calendar schema are running, but the calendar schema is not making any behavior proposals. The card schema’s behavior proposals require the focus stack, so the card-playing goal stays on the top of the stack. Then, triggered by the clock time, the calendar schema proposes a speech behavior about an upcoming appointment that requires the focus stack. Because the calendar schema has a higher priority, the arbitrator gives it control of the focus stack, which causes the calendar reminder goal to be pushed on top of the card-playing goal. The calendar reminder goal remains on the top of the stack throughout the (sub) dialogue regarding the appointment. When this reminder dialogue is completed, the calendar schema stops making proposals, the calendar goal is popped, and the arbitrator gives the focus stack resource back to the card schema, which has been continuously proposing the next behavior in the game, but never getting the needed focus stack resource.

Case 8) Changing Topic: Changing topic is handled similarly to Case 7, except that instead of the calendar schema deciding to make the interruption, the decision that the new topic has a higher priority than the current topic is made by some other schema based on cognitive reasoning that is outside of the scope of DiscoRT. The mechanism of pushing and popping the focus stack is identical, however. Furthermore, the interrupted schema may stop proposing the old topic, in which case it is never returned to.

Case 9) Barge-In: Barge-in is handled in the speech behavior realizer, which in addition to controlling the text-to-speech engine (voice resource), listens for events from the menu or speech perceptor (depending on the type of system). When such an event is received, the realizer immediately stops the text-to-speech engine and terminates.

5.3 Software Engineering and Privacy Concerns

Creating the working AlwaysOn system with DiscoRT and with twelve different user activities (plugins), some with direct manipulation interfaces, and others with on-screen presentation, required a significant testing effort. In addition, two plugins, one for engagement and one for the session plan also needed testing. Because the team intended for both agents to be available for up to a month of use, reliability of the entire system was a major concern. Each individual plugin, corresponding to one of the activities, was extensively tested in isolation. When all the plugins were integrated with the DiscoRT system, and the virtual agent or robot was introduced as user facing components, along with face recognition software and, in the case of the robot, control programs for its gestures, the final system required another round of testing, which took approximately four person-months to accomplish. Following that testing, the system with the virtual agent was beta-tested in the home of four elders after which more testing and debugging occurred.

Special requirements were needed to handle aspects of the robot agent. Because it made use of a camera in its eyes (rather than the one on the ASUS screen), adjustments to the face recognition algorithm to work with a moving robot head were undertaken. Control software to make it possible for the robot to nod at appropriate moments was introduced for the robot as well.

Because the users in the initial studies indicated substantial concerns about privacy, all of the data captured by the onboard and on-robot cameras in the study reported below was thrown away as soon as it was used by the face recognition software. The only video retained was that used in personal life story acquisition. All of the text of dialogs between users and the agents were gathered in text form, downloaded to a server, and stored behind a firewall for examination after study completion.

6 A STUDY OF THE AGENTS IN MONTH LONG TRIALS

To evaluate the efficacy of the AlwaysOn system, virtual agent and robot versions of the system were deployed into the homes of isolated older adults, where they interacted with the system daily for a month.

Following several other studies on comparisons between virtual and robotic agents, we hypothesized that the robot would be significantly more effective as a companion compared to the virtual agent. Our specific hypotheses were:

H1. Participants in the AlwaysOn robot condition will experience significantly greater decrease in loneliness compared to those in the AlwaysOn virtual agent condition, and those in the virtual agent condition will experience significantly greater decrease in loneliness compared to those in the Control condition, over the 30 day duration of the study.

H2. Participants in the AlwaysOn robot condition will experience significantly greater increase in happiness compared to those in the AlwaysOn virtual agent condition, and those in the virtual agent condition will experience significantly greater increase in happiness compared to those in the Control condition, over the 30 day duration of the study.

Experimental setup:

Recruitment for the study occurred over the course of one year through craigslist's posts, fliers, presentations at local community centers and mailed recruitment letters. These efforts resulted in seventy-five responses, of which forty-four were found eligible for the study. To be

eligible, participants need to be at least 55 years old, live alone, not exhibit major depressive symptoms (below a 3 on the PHQ-2 measurement given below) and be healthy enough to engage in basic physical activity (as assessed by the PAR-Q, given below). Following informed consent and screening, participants were randomly assigned to the robot, virtual agent, or control condition. Recruitment of participants took place over an entire year and at no time were there more than 4 participants with agents in their home. It was many months into the experiment before even a handful of participants' results were available for any of the conditions. Given the small numbers of participants assigned to each condition, results of the whole study were not available until a full year had passed.

In the robot and virtual agent condition, participants scheduled a time for the AlwaysOn system to be set up in their home. Once scheduled, a research assistant traveled to the participant's home to install the robot and/or touch screen computer. Participants chose where in their dwellings they wished to place the screen/robot. Almost all subjects choose their living rooms, with the exception of those who lived in studio apartments, in which case it was in the main room. Participants were then given a brief tutorial on how to use the system and were instructed to try and interact with it at least once a day for 30-days. No additional instruction was given to the participants. After the 30-day period passed, the research assistant returned to the participant's home to have them complete the debrief questionnaires and to collect the system.

Participants in the non-intervention control group were recruited in the same manner as other participants, and underwent the same intake procedure and baseline measurements. Following this, they were sent home and contacted 30 days later to return to the laboratory for final outcome measurements.

Study Measures:

During intake, the following questionnaires were given to participants to evaluate their health status and eligibility for the study:

- Sociodemographic Questionnaire (To assess age and relationship status)
- PAR-Q (To assess physical health status) (Thompson et al. 2013)
- PHQ-2 (To assess mental health status) (Löwe et al. 2005)

The following measures were given to participants at intake and debrief:

- Social Support Questionnaire (Norbeck et al. 1983)
- UCLA Loneliness Questionnaire (Russell and Peplau 1980)
- SF-12 Health Status (Bjorner and Turner-Bowker 2009)
- General Happiness (Lyubomirsky and Lepper 1999)

In addition, three questionnaires that assessed the participants' relation with a close friend or family member was given to assess any changes in the participant's relationship due to the system:

- ABC Questionnaire for Affective Benefits (Powell and Meyers 1995)
- Relationship Closeness Inventory (Berscheid et al. 1989)

Finally, the following two questionnaires, developed by the research team, were given during debrief, along with a semi-structured interview that focused on evaluating the system's efficacy and any memorable experiences they had with it:

- Working Alliance with Agent
- Social Agent Rating

Daily Measures:

In addition to the questionnaires given at intake and debrief, participants were given diary sheets to fill after each interaction with the system. These diary sheets consisted of four 5-point Likert

Table 1. Pre and Post Health Outcomes

Measures	Virtual Agent	Robot	Control	p
UCLA Loneliness	-0.09	-0.04	-0.1	0.9
Social Support	-1.12	-1.62	-0.67	0.55
General Happiness	0.21	0.25	0.58	0.55

Table 2. Overall Average System Usage

Measure	Virtual Agent	Robot	p
Total Sessions	30.5 (22.3)	27.5 (22.6)	.68
Total Time (mins)	248.5 (291.7)	241.0 (201.4)	.72
Days Used	22.2 (12.7)	21.5 (11.7)	.57

scale questions that assessed their mood and relationship with the agent. An open-ended section was also provided for participants to write any additional thoughts they had about the agent.

Results:

Forty-four people between the ages of 55 and 91 (Mean: 66, SD: 7.89) were recruited and randomized to one of three study conditions. This division resulted in eleven control participants, twenty-four virtual agent participants and nine in the robot condition. Out of these forty-four participants, thirty-six participants, ten in the control condition, eighteen in the agent condition and eight in the robot condition, completed the study. Of those who dropped out, one was due to a hardware failure with the robot, and seven dropped out voluntarily due to extraneous circumstances (e.g. ill family members or unexpected travel). Two participants dropped out due to annoyances with the system, related to the system’s brightness and noise.

The reason for the imbalance in randomization is that the study team originally had two virtual agent conditions to study the effect of different virtual agent designs. However, the team found no significant differences between the virtual agent groups, so combined them for all remaining analyses. Although the numbers of participants are different in each resulting condition, Levene’s tests demonstrated no significant differences in homogeneity of variances, supporting use of the one-way ANOVA tests used.

Health/Relationship Results:

As shown in Table 1, no significant changes were found in the participant’s health or relationship status across the three conditions. One way ANOVA tests were used for these results.

System Usage:

As presented in Table 2, no significant differences were found in overall system usage between the robot and the virtual agent. Non-parametric Mann-Whitney tests were used for system usage measures.

Further analysis of system usage, shown in Table 3 found that across all the conditions, **weather chat** was the most used module, followed by the **anecdotes** the agent told. Conversely, **video calls** were the least used module (no users tried it), followed by telling **stories** to the agent and the agent telling **about** itself.

Agent Ratings:

Tables 4 and 5 indicate that no significant differences between study condition and agent ratings. Non-parametric Kruskal-Wallis tests were used for this determination. The desire to have a conversation with the agent and how trustworthy it seemed were near significance, however (0.08), both of which favoring the robot condition.

Table 3. Average Usage per Activity

Module	Overall		Virtual Agent		Robot	
	Sessions	%Time	Sessions	%Time	Sessions	%Time
Enrollment	3.5	.04	3.1	.04	3.1	.05
About	3	.02	1.5	.02	1.5	.01
Explain	8.6	.12	8.5	.15	8.5	.06
Calendar	6	.04	6.5	.04	6.5	.04
Weather	25.1	.09	24.9	.07	25.5	13.6
Nutrition	5.9	.05	4.6	.05	4.6	.04
Exercise	8.9	.02	5.6	.02	5.6	.02
Health	4.7	.04	4.4	.04	4.4	.04
TicTacToe	8.7	.11	7.9	.12	7.9	.08
Checkers	6.5	.29	5.3	.29	5.3	.29
Story	1.3	.01	1.1	.01	1.1	.01
Anecdotes	21.4	.16	26.3	.13	26.3	.23

Qualitative Results:

Debrief interviews with the participant were recorded and transcribed for use in a categorical analysis in which themes related to companionship, social support, entertainment, realism and isolation were identified.

Companionship and Social Support:

During the debrief interviews, multiple participants remarked that the agent provided social support and companionship to them during their month-long interaction, claiming that the agent's personality allowed it to be "supportive but not judgmental. And humans can be judgmental even when they do not mean to be sometimes. Your body language sometimes comes across as eww. Not Karen" (P2). Multiple participants also found that the agent "kept [them] company" (P31), since it provided them with stories on demand, particularly for when there was "nobody I could talk to at 3 in the morning. And [the agent] was telling me stories about her friends when mountain climbing" (P4). However, they did find having her their home disruptive at times, since they often felt like they "had to meet with her every day. It was an inconvenience" (P19).

Entertainment:

In line with the recorded metrics, many participants reported that the anecdotal stories told by the agent were highly engaging, with participants feeling that "the stories, the jokes and the recipes were really fun. She could be the life of the party." The built-in games were also highly regarded, since the participants found it useful for keeping their "mind occupied because as you get older..."

Table 4. Agent and Robot Ratings

Question	Anchor 1	Anchor 7	Virtual Agent	Robot	p
How close do you feel to Karen?	Not at all close	Very close	3.62 (1.98)	3.71 (1.97)	0.64
How satisfied are you with Karen?	Not at all satisfied	Very satisfied	4.06 (1.78)	5 (1.29)	0.17
How much would you like to continue working with Karen?	Not at all	Very much	3.18 (1.88)	3.14 (2.04)	0.97
How much do you trust Karen?	Not at all	Very much	3.88 (1.73)	5.14 (1.35)	0.08
How much do you like Karen?	Not at all	Very much	4.53 (1.5)	5.14 (1.35)	0.35
Was Karen repetitive?	Not at all repetitive	Very repetitive	5.82 (1.47)	5.29 (1.38)	0.41
How easy was talking to Karen?	Very easy	very difficult	3.31 (1.89)	4 (2.16)	0.48
How interesting was Karen?	Not boring	Very interesting	3.81 (1.94)	4.71 (1.98)	0.33
How would you characterize your relationship with Karen?	Complete Stranger	Close friend	3.81 (1.47)	4.14 (1.78)	0.67
How much do you feel that Karen cares about you?	Not at all	Very much	3.88 (1.86)	3.43 (1.81)	0.6
How much do you feel that you and Karen understand each other?	Not at all	Very much	3.56 (1.86)	3.71 (2.14)	.87

Table 5. Agent and Robot Ratings (continued)

Question	Anchor 1	Anchor 7	Virtual Agent	Robot	p
How much do you feel that Karen was honest about her feelings towards you?	Not honest	Very honest	3.94 (1.81)	4.57 (1.33)	0.32
Karen behaved intelligently when we interacted.	Disagree completely	Agree completely	4.35 (2.15)	4.86 (1.77)	0.56
Karen understood our conversations.	Disagree completely	Agree completely	4.12 (1.76)	5.14 (1.35)	0.14
I wanted to have a conversation with Karen but forgot to do so.	Multiple times a day	Never	4.38 (1.73)	5.43 (0.98)	0.08

(P31). However, due to the nature of the provided games and their increased usage, participants seem to quickly exhaust them, mentioning that they wished there were “More games.” (P14) and that they should be “More challenging games. I would win every time and you want it to be a little bit of a challenge no matter how old you are; everyone seems to love game.” (P33).

Realism:

While a small set of participants found that interacting with the system “felt like talking to a person.” (P1), many participants were put off by the robotic nature of the agent’s voice and the repetitive options in her dialogue. Participants reported feeling like they “would miss her more if she had a better voice.” (P14) and that she “does not have a personality.” (P19).

Isolation:

The topics of isolation and the target population for the system came up multiple times during the debrief interviews, although they were often used as a way in which the participant distanced themselves from the rest of the older adult population. While most participants stated that they enjoyed using the system, they felt that it would be “more useful for a homebound senior” (P30) and that it was “more geared towards people [their] mother’s age.” (P42). These statements were made across a large range of ages, suggesting that these statements may be more indicative of the type of participants recruited for the study rather than the target age range.

6.1 Discussion of the Study

Our results do not support either of our study hypotheses, namely, that:

H1. Participants in the AlwaysOn robot condition will experience significantly greater decrease in loneliness compared to those in the AlwaysOn virtual agent condition, and those in the virtual agent condition will experience significantly greater decrease in loneliness compared to those in the Control condition, over the 30 day duration of the study.

H2. Participants in the AlwaysOn robot condition will experience significantly greater increase in happiness compared to those in the AlwaysOn virtual agent condition, and those in the virtual agent condition will experience significantly greater increase in happiness compared to those in the Control condition, over the 30 day duration of the study.

For two trending significant measures in the agent ratings, it is notable that trusting the robot more is in keeping with previous studies discussed earlier in this paper. Furthermore, the less human appearance of the Reeti robot agent may have contributed to users’ sense of trust because of the uncanny valley effect in creating robots that seem more human-like. Reasons for the second measure, that users wanted to have conversations more with the robot than the agent, are less clear. One possible explanation is that the physical form of the agents matters in the interactions. The robot’s physical presence and more apparent face tracking motions may have contributed to the desire for conversation. One could say that the robot seemed more “real” than the virtual agent. While both the robot and the virtual agent performed face tracking, this behavior was much more noticeable in the robot agent. However, whether these two differences or something else is at play is a matter for future study.

The only two values that approach trending, but are not low enough to be considered trends are: satisfaction with Karen (.17) and Karen understanding the conversation (.14). Interestingly in both cases, the Likert scores given by the participants are higher overall for both agents than any of the scores where there is no significance. Higher scores in these two questions can be seen as indicating that: 1) interacting with either agent was not a chore for the participants, 2) the agents were capable of enough conversational interaction to create a satisfying interaction and 3) the users liked both agents at a rating of between 4.5 and 5.1 on average. One can also note that users on average reported that the agents were not perceived as strangers, with a rating between 3.8 and 4.1.

At the same time, higher scores were also given to the claim that Karen was repetitive, which were rated above 5 on average by all participants. In anticipation of this problem, the AlwaysOn team attempted to design the conversations to limit repetition, which clearly did not succeed. Reducing repetitiveness in agents will continue to plague interaction that extends over long spans of time. More effort on dialog variability as well as techniques for agent personality and awareness of user emotion might address this challenging problem.

On the whole, the study participants had a positive, though not strongly positive attitude toward both agents, virtual and robotic. Furthermore, they interacted with the agent often and over many of the activities made available to them. For the extensive effort invested in creating this technology, this observation is satisfying to the AlwaysOn team. The reliability of the technology may have also contributed to the positive values of the study participants although no direct measures were made about participants' perceptions of reliability.

The reliability of the AlwaysOn agents is noted here. Based on culling the log files, there were 5 reliability issues that required in-person visits in the robot condition only. There were 9 reliability issues, 3 in the robot condition and 6 in the agent condition, that were fixed remotely. Considering the total number of days used by all users (approximately 572 across robot and virtual agent conditions) and the total of number of minutes by all users (approximately 4482 in the agent condition and 1928 in the robot condition), the team felt gratified by the operation of the AlwaysOn agents and the AlwaysOn system.

The research team believes that a valuable take-home lesson of the AlwaysOn project is the difficulty of getting participants for a study. Isolated adults are hard to reach by their very nature. A year of effort in tracking down participants by various methods did not yield as many participants as one would have liked, nor ones that were as isolated as statistics indicate exist in the US. Perhaps teaming with a hospital or large clinic would be a viable path in the future, although there are many challenges in patient privacy to be considered in such an undertaking.

7 CONCLUSIONS

This paper has presented the technology developed in the AlwaysOn project to provide a wide range of activities to a particular user population, isolated older adults. The interfaces for the technology included a cartoon-style virtual agent and a Reeti robot, both of which conversed with a user via speech output and menu input on a touch screen. They talked with users about activities designed to provide companionship and to reduce isolation through digital connections to friends and family and to the physical community. The technology provided a means to track user's faces, permit users to barge in during agent speech, provide for the agent to turn its face to an activity displayed onscreen, as well as to initiate interactions upon perceiving the user's face, but with awareness of the user's desires about whether to converse at any given time.

This work also presented early investigations on how the technology might be used, followed by the results of a month-long study in the homes of elder users. There are very few longitudinal studies of agent technology for isolated elders of more than a few participants, and in this regard, the effort was successful. While the results are not as conclusive as the team had hoped for, they do indicate that each user used the agents regularly and tried out many, though not all, of the activities developed. Trending results suggest that users will be more trusting of robots and more desirous of conversations with them. Overall, users indicated that conversations with both agents were satisfying to all the users, and users liked them. Also, not surprising given previous research, users indicated that the agents were not complete strangers to them.

The activities provided for the agents in AlwaysOn were pre-determined, and introduced by a simple set of rules. For longer term use, over many months or years, one could imagine a less

pre-determined method for introducing activities as well as introducing activities that were created after the system was being used, and dropping old ones that were not used. Methods might include learning from a group of users about which activities they first choose to use, or small chats that would focus on a single as yet unused activity to introduce it to the user. Such methods might alleviate the failure to use activities that the AlwaysOn project experienced.

One should not conclude that there is no place for virtual agents in home technology given the preference that seems to hold for robotic agents. It is quite conceivable that robots will be useful in the home, not only for the types of activities created for the AlwaysOn project, but when they can reliably move around and provide physical services to seniors. At the same time, virtual agents, perhaps looking just like the robot, will be useful on cell phones to help when a senior is out getting exercise, doing grocery shopping or providing a list of things to remember to ask one's physician about. Thus these technologies could form a continuum of use in the lives of seniors and non-seniors.

Finally, some comments about research technology in the home. To place technology for more than a few days in a user's home, elder or not, requires a significant commitment to testing, debugging and beta testing in order to assure the technology is reliable and bug proof. This level of effort is not normally associated with research projects. However, reliable technology is essential to understanding how technology plays a role in users' lives.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant Numbers IIS-0811942 and IIS-1012083. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

We acknowledge the programming and evaluation efforts of a group of students and staff: Dr. Barbara Barry, Elizabeth Bates, Morteza Behrooz, Bridgette Collado, Will Coon, Teryn Falkingham, Ayesha Fathima, Hannah Doolittle, Ashley Kline, JJ Liu, Kenneth Manning, Ramesh Manuvinakurike, Elise Masson, Karlyle Pilonis, Kathleen Totzke, John Connor Westfall, Mitchell Wills, and Zhe Zhang.

REFERENCES

- M. Argyle and M. Cook. 1976. *Gaze and Mutual Gaze*. Cambridge University Press. DOI: <https://doi.org/10.1017/S0033291700018523>.
- M. Behrooz, C. Rich and C. Sidner. 2014. On the Sociability of a Game-Playing Agent: A Software Framework and Empirical Study. *Proceedings of the Fourteenth International Conference on Intelligent Virtual Agents*. Springer, Switzerland. 40-53. DOI: https://doi.org/10.1007/978-3-319-09767-1_6.
- E. Berscheid, M. Snyder and A.M. Omoto. 1989. The Relationship Closeness Inventory: Assessing the closeness of interpersonal relationships. *Journal of personality and Social Psychology*, 57(5), 792. DOI: <https://doi.org/10.1037//0022-3514.57.5.792>.
- T. Bickmore, L. Caruso, K. Clough-Gorr and T. Heeren. 2005. "It's just like you talk to a friend" – Relational agents for older adults. *Interacting with Computers* 17(6) (December 2005), 711-735. DOI: <https://doi.org/10.1016/j.intcom.2005.09.002>.
- T. Bickmore, D. Schulman and L. Yin. 2010. Maintaining Engagement in Long-term Interventions with Relational Agents. *International Journal of Applied Artificial Intelligence*, 24(6) (July 2010), 648-666. DOI: <https://doi.org/10.1080/08839514.2010.492259>.
- T. Bickmore, R. Silliman, K. Nelson, D. Cheng, M. Winter, L. Henaulat, M. Paasche-Orlow. 2013. A Randomized Controlled Trial of an Automated Exercise Coach for Older Adults. *Journal of the American Geriatrics Society*, 61, 10 (September 2013), 1676-1683. DOI: <https://doi.org/10.1111/jgs.12449>.
- Jakob B. Bjorner and Diane M. Turner-Bowker. 2009. SF-36 and SF-12 Health Surveys. In: *Encyclopedia of Medical Decision Making*. Edited by: Michael W. Kattan. Volume 2, 1030-1036. DOI: <http://dx.doi.org/10.4135/9781412971980.n299>.
- Robert N. Butler. 1964. The life review: An interpretation of reminiscence in the aged. In R. Kastenbaum (ed.) *New Thoughts on Old Age*. Springer, Berlin Heidelberg. (1964), 65-70. DOI: https://doi.org/10.1007/978-3-662-38534-0_20.

- C. Chao, C. Lee, M. Begum, and A. Thomaz. 2011. Simon plays Simon says: The timing of turn-taking in an imitation game. *Proceedings of the Int. Symp. on Robot and Human Interactive Communication (RO-MAN 2011)*. IEEE Press, New York, NY, 235-240. DOI: <https://doi.org/0.1109/ROMAN.2011.6005239>.
- Juan Fasola and Maja J. Matarić. 2013. A Socially Assistive Robot Exercise Coach for the Elderly. *Journal of Human-Robot Interaction*, 2 (2) (2013) 3-32. DOI: <https://doi.org/10.5898/JHRI.2.2.Fasola>.
- J. Fodor and M. Roubens. 1994. *Fuzzy Preference Modeling and Multicriteria Decision Support*. Springer, The Netherlands, 1994. DOI: [10.1007/978-94-017-1648-2_8](https://doi.org/10.1007/978-94-017-1648-2_8).
- Maartje M.A. de Graaf, Somaya Ben Allouch and Jan A.G.M. van Dijk. 2017. Long-term evaluation of a social robot in real homes. *Interaction Studies: New frontiers in human-robot interaction*, 17, 3 (2016), John Benjamins Publishing Co. pp. 461-490. DOI: [10.1075/is.17.3](https://doi.org/10.1075/is.17.3)
- M. Granata, A. Chetouani, A. Tapus, P. Bidaud, and A. Dupourque. 2010. Voice and graphical-based interfaces for interaction with a robot dedicated to elderly and people with cognitive disorders. *Proceedings of the 19th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE New York, 785 - 790. DOI: <https://doi.org/10.1109/ROMAN.2010.5598698>.
- Erico Guizzo. Robosoft Unveils Kompai Robot To Assist Elderly, Disabled. *IEEE Spectrum Online*, March 9, 2010, retrieved February 1, 2017 from <http://spectrum.ieee.org/automaton/robotics/medical-robots/robosoft-kompai-robot-assist-elderly-disabled>.
- S. Hanke, E. Sandner, C. Tsiourti, and A. Stainer-Hochgatterer. 2015. The technical specification and architecture of a virtual support partner. *Workshop on Affective Interaction with Avatars, In conjunction with European Conference on Ambient Intelligence, Aml'15, Athens, Greece, 2015*.
- D. Hasegawa, J. Cassell, and K. Araki. 2010. The role of embodiment and perspective in direction-giving systems. *Dialog with robots: AAAI Fall Symposium*. 2010. Retrieved October 25, 2017 from <https://www.aaai.org/ocs/index.php/FSS/FSS10/paper/view/2186>.
- A. Holroyd, C. Rich, C. Sidner, and B. Ponsler. 2011. Generating connection events for human-robot collaboration. *Proceedings of the 20th IEEE Int. Symp. on Robot and Human Interactive Communication (ROMAN 11)*, IEEE Press, New York, 241-246. DOI: <https://doi.org/10.1109/ROMAN.2011.6005245>
- A. Holroyd and C. Rich. (2012) Using the behavior markup language for human-robot interaction. *Proc. ACM Conf. on Human-Robot Interaction*, ACM Press, New York, 2012, 147-148. DOI: [1 https://doi.org/10.1145/2157689.2157728](https://doi.org/10.1145/2157689.2157728)
- C. Huijnen, A. Badii, H. van den Heuvel, P. Caleb-Solly, and D. Thiemert. 2011. Maybe It Becomes a Buddy, But Do Not Call It a Robot - Seamless Cooperation between Companion Robotics and Smart Homes. *Proceedings of the Second International Joint Conference on Ambient Intelligence (AmI 2011)*. Springer, New York, 324-329. DOI: <https://doi.org/10.1007/978-3-642-25167-2>.
- D. O. Johnson, R. H. Cuijpers, J. F. Juola, E. Torta, M. Simonov, A. Frisiello, M. Bazzani, W. Yan, C. Weber, S. Wermter, N. Meins, J. Oberzaucher, P. Panek, G. Edelmayer, P. Mayer and C. Beck. 2014. Socially Assistive Robots: A Comprehensive Approach to Extending Independent Living. *International Journal of Social Robotics*. 6 (2) (April 2014). 195-211. DOI: <https://doi.org/10.1007/s12369-013-0217-8>
- Adam Kendon. 1967. Some functions of gaze direction in two person interaction. *Acta Psychologica*, 26 (1967), 22-63. DOI: [https://doi.org/10.1016/0001-6918\(67\)90005-4](https://doi.org/10.1016/0001-6918(67)90005-4).
- C. Kidd and C. Breazeal. 2004. Effect of a Robot on User Perceptions. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004)*, IEEE Press, New York, NY, 3559 - 3564 vol.4. DOI: <https://doi.org/10.1109/IROS.2004.1389967>.
- C. D. Kidd and C. Breazeal. 2008. Robots at Home: Understanding Long-Term Human-Robot Interaction. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2008)*, IEEE Press, New York, NY, 3230 - 3235. DOI: <https://doi.org/10.1109/IROS.2008.4651113>.
- T. Klamer, S. Ben Allouch, and D. Heylen. 2010. "Adventures of Harvey" - Use, Acceptance of and Relationship Building with a Social Robot in a Domestic Environment. *Third International Conference, (HRPR 2010)*, Springer, Berlin, 74-82. DOI: https://doi.org/10.1007/978-3-642-19385-9_10.
- B. Löwe, K. Kroenke, and K. Gräfe. 2005. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *Journal of Psychosomatic Research*, 58(2), 163-171. DOI: <https://doi.org/10.1016/j.jpsychores.2004.09.006>.
- S. Lyubomirsky and H. Lepper. 1999. A measure of subjective happiness: preliminary reliability and construct validation. *Soc. Indic. Res.* 46, 137-155. DOI: <https://doi.org/10.1023/A:1006824100041>.
- D. Maciuszek, J. Åberg, N. Shahmehri. 2005. Evaluation and refinement of a design framework for generating dependable virtual companions for later life. *From Smart Homes to Smart Care: 3rd International Conference on Smart homes and health telematics, (ICOST 2005)*, IOS Press, Amsterdam, The Netherlands, 50-64. Retrieved October 25, 2017 from https://books.google.com/books?hl=en&lr=&id=7AWMY0uz8awC&oi=fnd&pg=PA50&ots=gDnb8qJiCm&sig=2qcfZqslJp9whcOv-x-q_-EocQ#v=onepage&q&f=false.

- Marcus Mast, Michael Burmester, Katja Krüger, Sascha Fatikow, Georg Arbeiter, Birgit Graf, Gernot Kronreif, Lucia Pigini, David Facal and Renxi Qiu. 2012. User-Centered Design of a Dynamic-Autonomy Remote Interaction Concept for Manipulation-Capable Robots to Assist Elderly People in the Home. *Journal of Human Robot Interaction*, Vol 1(1), 2012. DOI: <https://doi.org/10.5898/jhri.1.1.mast>.
- Miraculous Life. 2013. Retrieved February 1, 2017 from <http://www.miraculous-life.eu/index.php>.
- M. Nani, P. Caleb-Solly, S. Dogramadgi, C. Fear and H. van den Heuvel. 2010. MOBISERV: An Integrated Intelligent Home Environment for the Provision of Health, Nutrition and Mobility Services to the Elderly. 4th Companion Robotics Workshop, Brussels, Belgium, 2010. Retrieved October 25, 2017 from https://www.researchgate.net/publication/268346147_MOBISERV_An_Integrated_Intelligent_Home_Environment_for_the_Provision_of_Health_Nutrition_and_Mobility_Services_to_the_Elderly.
- N. R. Nicholson. A Review of Social Isolation: An Important but Underassessed Condition in Older Adults. *The Journal of Primary Prevention*, 33, 2-3, (June 2012), 137-152. DOI: <https://doi.org/10.1007/s10935-012-0271-2>.
- J.S. Norbeck, A.M. Lindsey, and V.L. Carrieri. 1983. Further development of the Norbeck Social Support Questionnaire: Normative data and validity testing. *Nursing Research*, 32, 4-9. DOI: <https://doi.org/10.1097/00006199-198301000-00002>.
- B. Nooraei, C. Rich and C. Sidner. 2014. A Real-Time Architecture for Embodied Conversational Agents: Beyond Turn-Taking. 7th International Conference on Advances in Computer-Human Interactions, Barcelona, Spain, May, 2014. Retrieved October 25, 2017 from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.700.6699>.
- Hugo Pinto, Yorick Wilks, Roberta Catizone and Alexiei Dingli. 2008. The senior companion multiagent dialogue system. *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems (May 2008)*, 1245-1248. Retrieved from <https://dl.acm.org/citation.cfm?id=1402842>.
- L.E. Powell and A.M. Myers. 1995. The Activities-specific Balance Confidence (ABC) Scale. *J Gerontol A Biol Sci Med Sci*. 1995;50A: M28-34. pmid:7814786. DOI: <https://doi.org/10.1037/t13026-000>.
- M. Reblin and B.N. Uchino. 2008. Social and emotional support and its implication for health. *Current Opinion in Psychiatry*, 21, 2 (March 2008), 201-205. DOI: <https://doi.org/10.1097/YCO.0b013e3282f3ad89>
- Reeti, a smart compagnon, a multimedia tool. 2017. Retrieved March 1, 2017 from <http://www.reeti.fr/index.php/en/detailen>.
- P.H. Robert, A. Konig, H. Amieva, S. Andrieu, F. Bremond, R. Bullock, M. Ceccaldi, B. Dubois, S. Gauthier, P.A. Kenisgberg, S. Nave, J.M. Orgogozo, J. Piano, M. Benoit, J. Touchon, B. Vellas, J. Yesavage and V. Manera. Recommendations for the use of Serious Games in people with Alzheimer's Disease, related disorders and frailty. *Frontiers in Aging Neuroscience*. 2014; 6(54), (March 24, 2014). DOI: [10.3389/fnagi.2014.00054](https://doi.org/10.3389/fnagi.2014.00054).
- L. Ring, L. Shi, K. Totzke, T. Bickmore. Social Support Agents for Older Adults: Longitudinal Affective Computing in the Home. *Journal on Multimodal User Interfaces*. 2015;9(1):79 -88. DOI: <https://doi.org/10.1007/s12193-014-0157-0>.
- C. Rich and C. Sidner. 1999. Collagen: A collaboration manager for software interface agents. *User Modeling and User-Adapted Interaction*, 8, 3/4, (September 1998), 315-350. DOI: [10.1023/A:1008204020038](https://doi.org/10.1023/A:1008204020038).
- C. Rich, C. Sidner and N. Lesh. 2001. Collagen: Applying collaborative discourse theory to human-computer interaction. *AI Magazine*, 22, 4, (Winter 2001), 15-25. DOI: <https://doi.org/10.1609/aimag.v22i4.1589>
- C. Rich. 2009. Building task-based user interfaces with ANSI/CEA-1818. *IEEE Computer*, 42, 8, (August 2009), 20-27. DOI: <https://doi.org/10.1109/MC.2009.247>.
- C. Rich, B. Ponsler, A. Holroyd, and C. Sidner, 2010. Recognizing engagement in human-robot interaction. *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*. IEEE Press, Piscataway, NJ. 375-382. DOI: <https://doi.org/10.1109/HRI.2010.5453163>.
- C. Rich and C.L. Sidner. 2012. Using collaborative discourse theory to partially automate dialogue tree authoring. In: *Proc. 12th Int. Conf. on Intelligent Virtual Agents (IVA 2012)*, Springer, Berlin, 327-340. DOI: https://doi.org/10.1007/978-3-642-33197-8_34.
- D. Russell and L. Peplau. 1980. The revised UCLA loneliness scale: concurrent and discriminant validity evidence. *Journal of Personality and Social Psychology*. 39, (September, 1980), 472-80. DOI: <https://doi.org/10.1037/0022-3514.39.3.472>.
- S. Sabanovic, T. Bennett, W. Chang, and L. Huber. 2013. PARO robot affects diverse interaction modalities in group sensory therapy for older adults with dementia. *IEEE International Conference on Rehabilitation Robotics (ICORR 2013)*. IEEE Press, Piscataway, NY. 1-6. DOI: [10.1109/ICORR.2013.6650427](https://doi.org/10.1109/ICORR.2013.6650427).
- Emanuel A. Schegloff and Harvey Sacks. 1973. Opening Up Closings. *Semiotica*. 8, 4, (November 2009). 289-327. DOI: <https://doi.org/10.1515/semi.1973.8.4.289>.
- A. Seiderer, S. Hammer, E. Andre, M. Mayr and T. Rist. 2015. Exploring Digital Image Frames for Lifestyle Intervention to Improve Well-being of Older Adults. *Proceedings of the 5th International Conference on Digital Health (DH 15)*. ACM New York, NY, 71-78. DOI: <https://doi.org/10.1145/2750511.2750514>.
- C.L. Sidner, C. Lee, C. Kidd, N. Lesh, and C. Rich. 2005. Explorations in engagement for humans and robots. *Artificial Intelligence*. 166, 1-2, (August 2005), 140-164. DOI: <https://doi.org/10.1016/j.artint.2005.03.005>.

- C.L. Sidner, C. Rich, M. Shayganfar, T. Bickmore, L. Ring and Z. Zhang. 2015. A Robotic Companion for Social Support of Isolated Older Adults. In *HRI'15 Extended Abstracts Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*, ACM Press, New York, NY, 289. DOI: <https://doi.org/10.1145/2701973.2702103>. Video at: <https://www.youtube.com/watch?v=HngGcvEuB60>, retrieved January 1, 2017.
- Andrew Steptoe, Aparna Shankar, Panayotes Demakakos, and Jane Wardle. Social isolation, loneliness and all-cause mortality in older men and women. *Proceedings of the National Academy of Sciences of the USA*. 110, 15, (April 9, 2013), 5797-5801. DOI: <https://doi.org/10.1073/pnas.1219686110>.
- Paul D. Thompson, Ross Arena, Deborah Riebe and Linda Pescatello. 2013. ACSMs New Preparticipation Health Screening Recommendations from ACSMs Guidelines for Exercise Testing and Prescription, Ninth Edition. *J Current Sports Medicine Reports*. 12(4). (July/August 2013) 215-217. DOI: <https://doi.org/10.1249/jsr.0b013e31829a68cf>.
- Christiana Tsiourti, Emilie Joly, Cindy Wings and Katarzyna Wac. Virtual Assistive Companions for Older Adults: Qualitative Field Study and Design Implication. Conference: *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare (May 2014)*. 57-64. DOI: 10.4108/icst.pervasivehealth.2014.254943.
- L. Pfeifer Vardoulakis, L. Ring, B. Barry, C. Sidner and T. Bickmore. 2012. Designing Relational Agents as Long Term Social Companions for Older Adults. 12th International Conference on Intelligent Virtual Agents (IVA 2012). Springer Verlag, Berlin, 289-302. DOI: <https://doi.org/10.1007/978-3-642-33197-8>.
- H. Vilhjálmsón, N. Cantelmo, J. Cassell, N. E. Chafai, M. Kipp, S. Kopp, M. Mancini, S. Marsella, A.N. Marshall, C. Pelachaud, Z. Ruttkay, K.R. Thórisson, H. van Welbergen, R. J. van der Werf. 2007. The Behavior Markup Language: Recent Developments and Challenges. In: C. Pelachaud, J.C. Martin, E. André, G. Chollet, K. Karpouzis, D. Pelé (eds). *Intelligent Virtual Agents. IVA 2007. Lecture Notes in Computer Science*, vol 4722. Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-540-74997-4_10.
- K. Wada, T. Shibata, T. Saito and K. Tanie. 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 International Conference on Intelligent Robots and Systems (IROS 2003)*, 3, IEEE Press, 2847-2853. DOI: <https://doi.org/10.1109/IROS.2003.1249302>.
- J. Wainer, D. Feil-Seifer, D. A. Shell, and M. J. Matarić. 2007. Embodiment and Human-Robot Interaction: A Task-Based Perspective. *Proceedings, 16th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2007)*, IEEE Press (2007), 872-877. DOI: <https://doi.org/10.1109/ROMAN.2007.4415207>.
- K. Windle, F. Francis and C. Coomber. 2011. Preventing loneliness and social isolation: Interventions and outcomes. London: Social Care Institute for Excellence (October 2011), 16. Retrieved October 25, 2017 from <https://www.scie.org.uk/publications/briefings/briefing39/>.
- J. Wrobel, Y. Wu, H. Kerhervé, L. Kamali, A. Rigaud, C. Jost, B. Le Pévédic and D. Duhaut. 2013. Effect of Agent Embodiment on the Elder User Enjoyment of a Game. *The Sixth International Conference on Advances in Computer-Human Interactions (ACHI 2013)*. ThinkMindTMDigital Library, (February 2013), 162-167. Retrieved October 25, 2017 from <https://hal.archives-ouvertes.fr/hal-00832097/>.
- R. Yaghoobzadeh, M. Kramer, K. Pitsch, S. Kopp. 2013. Virtual Agents as Daily Assistants for Elderly or Cognitively Impaired People, Studies on Acceptance and Interaction Feasibility. *Intelligent Virtual Agents, 13th International Conference on Intelligent Virtual Agents (IVA 2013)*, Springer Verlag, Berlin, 79-91. DOI: <https://doi.org/10.1007/978-3-642-40415-3>.
- V. Yngve. 1970. On getting a word in edgewise. *Papers from the Sixth Regional Meeting of the Chicago Linguistics Society*, University of Chicago, Department of Linguistics, 567-578.

Received