CATE: An Embodied Conversational Agent for the Elderly

Sean Latrelle Bravo, Cedric Jose Herrera, Edward Carlo Valdez, Klint John Poliquit, Jennifer Ureta, Jocelynn Cu, Judith Azcarraga and Joanna Pauline Rivera

College of Computer Studies, De La Salle University, Manila, Philippines
{jennifer.ureta, sean.bravo}@dlsu.edu.ph

Keywords: Embodied Conversational Agent, Elderly Companionship.

Abstract: Social isolation and loneliness affect one-third to one-half of the elderly population, and these have a very negative impact on the physical and mental health of the elderly. Cate is an Embodied Conversational Agent designed to alleviate social isolation and loneliness of the elderly. It takes in user inputs through buttons and responds through text-to-speech and emotion animations developed with the use of SmartBody. Interviews were conducted to identify what the elders like to talk about and how elders interact in general. An end-user evaluation form and an observation checklist were used to see what the elders thought of the application. Cate achieved an overall satisfaction score of 3.98 on a scale of 1 (lowest) to 5 (highest). Based on the results, elders generally like interacting with Cate and can see Cate as a possible companion.

1 INTRODUCTION

Over the past two decades, the number of elderly people in the Philippines grew to 9.4 million constituting 8.6% of the population (ABS-CBN, 2018). Elderly people in the Philippines are traditionally cared by the family members or relatives. However, as many Filipinos try to find work abroad in an attempt to have a better paying job, their collective effort to care for the elderly dwindles (Antonio, 2015). Nowadays, traditional families are separated by distance, time, and/or lack of understanding (Jones, 2017).

One-third to one-half of the elderly population are affected by social isolation and loneliness. These have negative impact on the physical and mental health of the elderly. Social isolation have been identified as a risk factor for morbidity and mortality with outcomes comparable to smoking, obesity, and hypertension. It has also been linked to decrease in resistance to infection, cognitive decline, and mental health conditions such as depression and dementia (Landeiro et al., 2017).

To reduce social isolation and loneliness, two main types of interventions were identified by Landeiro et al (2017): group-based interventions (e.g., support groups, reminiscence therapy, and videoconferencing) and one-to-one interventions (e.g., computer use training, animal companionship, and visitor volunteers). These interventions can either be technology-assisted or not (Landeiro et al., 2017).

There are already numerous assistive technologies for elderly. Some of these are CareZone for medication (Carezone, 2018), and Elderly Care for motivation, inspiration, pharmaceutical store locating, and socializing (AB, 2017). Among these, only a few deals with the emotional state of elderly.

Embodied Conversational Agents (ECA) are conversational agents which can be in a form of a machine or a software application, and has a virtual representation of a human (Bevacqua et al., 2007). These can be designed to provide support and counseling to a person. There are several ECAs that are used mainly for conversing and interacting. Some of these are Ellie (DeVault et al., 2014), Rea (Bickmore and Cassell, 2005), Greta (Niewiadomski et al., 2009), and Justin/Justina (Kenny et al., 2008). However, very few ECAs cater to elderly people despite its potential as a virtual companion.

A conversational agent can be of help in alleviating social isolation and loneliness of elderly in three ways (Ring et al., 2013). First, to reduce their loneliness, an agent could serve as a companion and social support by engaging them to small talk and/or to social activities such as games (Bickmore and Cassell, 2005). Second, to reduce their social isolation, an agent can connect them to their friends and relatives through electronic communication, visit and chat coordination, and interventions in their social behavior. Third, an agent can perform a talk therapy (Colby, 1995).
This work focuses on developing a virtual companion for elderly people through the use of ECA to help alleviate their social isolation and loneliness. In this paper, we discuss the dialogue design of Cate which includes the dialogue types and the conversation topics derived from the interviews with the elderly. The other activities in the system are also described. Lastly, we show our evaluation metrics and analysis of results.

2 RELATED WORK

There are a few studies that focuses on providing social support for older adults using agents.

A preliminary study was conducted by Mival et al (2004). They made use of AIBO, a robotic dog, as a companion for elder adults during a chess game. They have concluded that the agent has to be proactive in starting a conversation. Users find an agent that initiates a conversation more desirable. Another conclusion is that utility is more important to the elderly than micro level details of the agent like physical characteristics, behaviours, and personality. Once its usefulness is established, users get engaged to the interaction and will start to notice the micro level details (Mival et al., 2004).

Another preliminary study was conducted by Vardoulakis et al (2012). They deployed an Embodied Conversational Agent (ECA) in a home for the elderly to be tested by 12 elderly. The ECA can be controlled remotely by the research assistants. The satisfaction rate of their system is 6 on a scale of 1 (lowest) to 7 (highest). From the interactions, they have identified conversation topics (e.g. storytelling, family, friends, music, news, fashion, wheather, activity planning, and attitude towards aging) and general design principles for an ECA for the elderly. The main limitation of their system is the in-home video recording that made several participants very uncomfortable. The in-home video recording was necessary since the research assistants need to see and hear the participants for them to respond appropriately. Thus, their recommendation is to create an autonomous agent that would not need a support of a human (Vardoulakis et al., 2012).

Following the aforementioned studies is the work of Ring et al (2013). They developed an ECA to provide automated social support for elderly. The user selects responses from multiple choices provided by the ECA. Then, the ECA responds by nodding its head, speaking, and/or gesturing. Its dialogue design has two components: (a) companionship and social support which assesses user’s affect state then provides empathetic feedback, and (b) loneliness and depressive symptom interventions which provides motivation and inspiration. The ECA was evaluated by 12 elderly. It has a satisfaction rate of 4.4 on a scale of 1 (lowest) to 7 (highest) and an ease-of-use rate of 1.9 on a scale of 1 (highest) to 7 (lowest). Their results suggest that their system is effective in assessing and managing user’s affect state, and that making the agent proactive in starting a conversation highly increased the effectiveness of the system. However, the agent being a social support needs to be improved by facilitating social connectivity (Ring et al., 2013).

3 CATE: CONVERSATIONAL AGENT FOR THE ELDERLY

Cate is an ECA that can converse with elderly users. She is embedded in a mobile application that is available on smartphones running Android 5.0 and above.

The user can communicate with Cate through buttons and text fields that are displayed on the screen. Cate, in turn, responds through SmartBody’s text-to-speech module. Depending on the choices of the user, the kind of dialogue is selected from the dialogue templates, and the appropriate emotion (i.e. happy, concerned, or worried) is expressed.

Cate’s age range from late twenties to early thirties. Her upper body is shown. Her eyes, eyebrows, mouth, neck, and hair are used to show Cate empathizing with the user. SmartBody, a character animation platform, is used for the facial animations. A woman’s voice, who is fluent in the English language, is used as the voice of Cate. Both the animation and the speech are synchronized when Cate’s response is delivered.

3.1 Dialogue Design

An interview was conducted at a house for the elderly. There were a total of three interviewees. During the interview with the elderly, they talked about family at some point. However, it seemed like a sensitive topic for them as some of them were wondering why they were left there by their family. This is the reason why family is not one of the preset topics, which are topics always suggested by Cate when the user wants to talk about something in the application. When the elderly were asked what they usually talk about, all of them said they liked talking about what they watched on the television in the house.
3.1.1 Dialogue Types

Cate’s responses are based on four dialogue types: immediate feedback (i.e. positive, negative, and neutral), pumping, suggestion, and summarizing. These dialogue types are used as templates for the generation of responses. Cate can provide positive, neutral, and negative feedback. She can also ask the user for more information (e.g. “What else?” or “What more can you tell me?”). She can also suggest some activities found in Section 3.2. Since the users’ inputs are mostly limited to buttons, dialogue types are usually pre-determined when a user makes a choice. An example of a dialogue with Cate can be seen at Figure 1.

Figure 1: A sample draft of conversation with Cate.

3.1.2 Conversation Topics

Cate has ten conversation topics: getting to know, feeling, jokes, food, reading, television, pets, cat or dog pictures, family, and social media. The first topic is getting to know. The preset topics are food and reading, while the random topics are feeling, jokes, television, pets, cat or dog pictures, family, and social media.

a. Getting to Know
Cate will ask for the user’s first and last name, birthday, gender, and hobby. Then, Cate will start a conversation by asking the user how they currently feel. The conversation topic will then be changed to feelings. Afterwards, Cate will ask if the user wants to continue talking. If the user agrees, Cate will suggest a different conversation topic, or an activity.

b. Feeling
Whenever Cate asks how the user is feeling, there are always 3 choices which will lead the conversation to three different paths: happy, sad, and angry. The happy path is where Cate encourages the user to continue whatever made them happy as shown in Figure 1. The sad path is where Cate tries to sympathize with the user, and tries to make them happy. The angry path is where Cate will try to calm down the user first, then will ask them if they want to talk about it. Cate will not pry as it might add more fuel to the fire, but will try her best to make them calm down and talk about the issue at hand.

c. Jokes
A good humor or joke may ease the struggles that they are going through in life. It may also help them forget about their sadness (Flores, 2018). Cate can tell them a joke to hopefully uplift their mood.

d. Food
The topics for food are about two of the famous fast-food chains in the Philippines. One is a long-time establishment, while the other is a relatively modern fast-food restaurant. These two restaurants represent the past and the present. Talking about the past gives the elderly people nostalgia making them feel better (Gergov and Stoyanova, 2013).

e. Reading
Scientific studies show that reading helps with enhancing memory and sharpening decision skills. It can also help reduce an elderly person’s stress to the point that they sleep better than usual (HomeInstead, nd).

Cate will start by asking what the user usually reads (i.e. books, magazines, or newspapers), then she will ask about the title and genre. If they choose books, she will also ask if they like fiction or non-fiction. Cate will also share her preferences to the user. Afterwards, she will recommend the user to read more and will motivate them by citing benefits of reading.

f. Television
The volunteers in the participating house for the elderly said that elderly love watching television (TV). Most of the time, the random topics they talk about with one another include what they recently watched on TV (Vardoulakis et al., 2012).

Cate will ask about the title of their favorite show, their purpose of watching TV (i.e. information, or entertainment), and their favorite TV channel. If they answered that they watch TV for information, they will be asked if they usually watch news or TV series. On the other hand, if they answered entertainment, they will be asked if they usually watch sports or TV series. Afterwards, Cate will ask how long they usually watch TV and will cite effects of watching TV.

g. Pets
Studies have shown that having pets have potential health improvements, both physical and men-
tual health, for the elderly (Cherniack and Cherniack, 2014).

Cate will ask about the user’s pet, the name of the pet, and their sentiment about their pet. She then asks how they got the pet. If the user says it was given a family or friend, Cate will suggest to the user get in touch, and thank them through phone.

**h. Cat or Dog Pictures**

Looking at animal pictures tends to make people happy due to their cuteness (Nittono et al., 2012). Thus, Cate will sometimes, show pictures of cute dogs and cats which are two of the most common domesticated pets.

**i. Family**

Elderly always talk about how their families take care of them and love them (Vardoulakis et al., 2012). For this topic, Cate will ask if the user talked to their family recently. If they do, she will ask about what they talked about and how they talked (i.e. thru phone or in person). Otherwise, she will encourage the user to talk to their family, and suggest the user to call them, or set a reminder when to call them.

**j. Social Media**

Social media strengthens ties between friends and family by making them engage with one another (Cornejo et al., 2013). When the users pick the social media topic, Cate will ask them if they use social media. If they do, Cate will ask them what social media platform they use and for what purpose do they use it for (i.e. social interaction, or news and events updates). If they use social media to interact with people, Cate will ask who they usually talk to (i.e. family or friends), how often, and will they meet soon. She will then suggest to the user to set a reminder to talk to their family or friends.

### 3.2 Activities

Cate features three different activities, aside from conversations, for the elderly users to help alleviate their social isolation and loneliness, or to assist them.

#### 3.2.1 Calls

Cate could suggest to make a call for the user. For example, at the end of the pets topic, the user is given a choice whether to call their family or friend who gifted them the pet. Another instance of this is in the getting to know path when the user picks that someone made them happy because of a gift. Figure 2 shows that Cate is suggesting to make a call as one of the other activities. Figure 3 shows that Cate is suggesting some contacts for the user to call to thank them through the phone.

#### 3.2.2 Reminder (Calendar)

Cate could suggest to the user to set a reminder. The application will open the calendar application of the phone, and the user will confirm the contents of the reminder. Once set, the Calendar application will close and will go back to Cate. For example, the user can set a reminder at the end of the family topic. If the user says that they are seeing their family soon, Cate asks if they want to set a reminder to when this meeting will happen. Same as the call, another instance is when the user is happy because someone gave them a gift. Figure 2 shows that Cate is suggesting to set a reminder as one of the activities. Figure 4 shows that Cate opened the calendar application of the phone for the user.

#### 3.2.3 Games

Cate will sometimes ask the user if they would want to do something else. If they do, Cate could suggest to play a game. There are currently two types of simple games in the application: trivia and picture game.

**a. Trivia**

Trivia is a quiz-like game where Cate will ask the user some questions about history because it is one of the most exciting trivia game topics for elderly (Super-Carers, nd). Sample questions are: “Who waved the flag during the Philippines independence proclamation in 1898?””, “Who coined the name of the longest-running noontime show “Eat Bulaga!”?”, and “Is
Manny Pacquiao’s nickname Pac-man based on a Japanese arcade game of the same name?”. b. Picture Game
The picture game shows the user some landmarks, buildings, and churches of the Philippines. It may give a sense of nostalgia since the landmarks and photos used were from the past, helping them reminisce about the good old days. Some of the sample photos used in the picture game are the old photos of the Malacañang Palace, Manila City Hall, and Araneta Coliseum (FilipiKnow, 2018).

3.3 Knowledge Base
In Cate’s knowledge base, since the dialogues are template-based, there are three major stored data: user profile, questions, and responses. The user profile contains the user’s first and last name, birthday, gender, and hobby. The name is asked so that Cate can keep mentioning their name when having their conversation to build a sense of companionship. Questions are what Cate can ask to the user. Responses are Cate’s reactions and suggestions regarding the user’s answer. It contains the identifier of the current decision, a string of decision identifiers that will consequently display the next question/response/dialogue, the choices that the user will be having next or the text-field that the user will be answering, and the tag of the emotion (i.e. happy, concerned, or worried) that Cate will be executing.

4  METHOD

4.1 Participants
Ten participants aged 60-77 years old were chosen to test the application. Four of them were from a retirement home and the other four were living with their family. The elderly from the retirement home were either abandoned by their family or were placed their by the government’s social welfare

For some of the elderly who had difficulties with understanding English, especially those from the retirement home, some of the conversations and evaluation questions had to be translated by the researchers to the local language. Some of the participants also had to be assisted because they were technologically challenged.

4.2 Privacy and Ethical Issues
Since Cate was evaluated by human participants, an informed consent form was given to them. For those who have disabilities, they are assisted by a family member, or the volunteers in the participating house for the elderly. The elderly were informed that their participation is voluntary, and that they may choose to withdraw anytime without any consequences. The system, testing procedure, potential risks and discomforts, potential benefits to subject or society, and confidentiality were discussed in the informed consent form. There might be a potential discomfort in dealing with an ECA since they are not really human, but only has avatar as its human representation. There are no potential risks or harm that may be posed to the participants. The information obtained during the study regarding the user was kept confidential and their identity was kept anonymous.

4.3 Evaluation
Cate was evaluated by the participants by answering an evaluation form which is split into 7 categories: naturalness, embodiment, interaction and affect, joy of use, ease of use, acceptability, and utility. For each category, the participants are asked to answer a set of questions using a Likert scale of 1 (strongly disagree) to 5 (strongly agree).

The Naturalness category is used to evaluate how human-like the virtual agent is when interacting with the elderly. Embodiment deals with evaluating the external appearance and personality of the virtual agent. It is also used to measure how well the agent is able to convey itself as likeable through its facial expression, speech and gestures. Interaction and affect deals
with how the elderly feel when they interact with Cate while using the application. It is used to evaluate how emphatic and comforting the agent is. Joy-of-use is a subcategory under the interaction and affect category. It refers to the positive feeling the user has when using the application. Ease-of-use is used to evaluate how much mental effort is required from the user to do a certain task and how user friendly it is. This is important because we have to consider the cognitive and physical abilities of the elderly. Acceptability is used to determine if the user likes using the application and if they would recommend it to other people. Utility is used to evaluate the functionality of the system related to the task that the user wants to accomplish. The agent must be capable of meeting the end goals of the user.

The researchers also answered an observation checklist while the users were talking to Cate, in order to get more insights and data to support the scores and answers in the end user evaluation.

5 RESULTS AND ANALYSIS

5.1 Overall Results

Table 1: This table shows the average score for each category in the end-user evaluation form.

<table>
<thead>
<tr>
<th>Category</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naturalness</td>
<td>3.8</td>
</tr>
<tr>
<td>Embodiment</td>
<td>3.73</td>
</tr>
<tr>
<td>Interaction and Affect</td>
<td>4.18</td>
</tr>
<tr>
<td>Joy-of-use</td>
<td>3.99</td>
</tr>
<tr>
<td>Ease-of-use</td>
<td>4.03</td>
</tr>
<tr>
<td>Acceptability</td>
<td>3.68</td>
</tr>
<tr>
<td>Utility</td>
<td>4.29</td>
</tr>
<tr>
<td>Total Mean</td>
<td>3.98</td>
</tr>
</tbody>
</table>

Table 1 shows the average results for each category. Utility had the highest score with an average of 4.29 while the lowest score was for acceptability. Under the naturalness category, the lowest scoring criteria was the “I can sense emotions through Cate’s voice”, which is under the naturalness category. It had an average score of 3 which is neutral. However, this is caused by a limitation of the system on how Cate pronounces certain words. This limitation made it difficult to determine the emotion in her speech for some people. Some also found the voice too robotic for their liking. “The conversation with Cate is natural” and “Talking to Cate is similar to feels like talking to a human being” both had the highest average score of 4.3. The elderly said that the conversations felt natural and that they loved talking to Cate because of how she would keep asking them questions. For embodiment, the criteria with the highest average score was “I like the way Cate looks” with an average score of 4.2. Based on the observation checklist and the evaluation scores, some were amazed by how human-like she looked while some thought she looked like a cartoon character. One of the participants initially remarked that Cate looked snobbish but she eventually like talking to her after a while, after seeing how she looked and acted throughout their conversation.

Interaction and affect had the second highest average score. The highest average scores under this category were from the criteria “Cate was friendly” and “Interaction with Cate is not repetitive”. Overall, most of the elderly liked talking to Cate. Although some of them did not like how she looked, they liked interacting with her. One elderly said that they were amazed by how she kept talking to them while another even began to verbally talk to Cate because she was so immersed in the conversation, that she forgot that Cate does not accept voice input. There was also a notable instance with one participant wherein she was asked by Cate about her first crush, first date and first kiss and she laughed and smiled and said to the researchers: “I forgot already, but it is nice to know that she is curious about me”. The lowest scores encountered were for the criteria “The random topics that Cate offered me were interesting” and “The jokes that Cate shared were funny”. The scores were 3.7 and 3.6 respectively. The reason for the low scores for the random topics could be because they got a topic they did not like. However, for some of them they really enjoyed the topics that were randomly assigned to them. One of the elderly participants whose favorite hobby was reading, was amazed when she saw reading as a topic and when Cate talked about it with her. For the jokes, the low scores could be due to Cate’s voice. An essential part of telling the joke is delivery, and Cate’s robot-like voice is not able to handle this well.

The category ease of use had a high score. The only problem that the participants had, especially those from the retirement home, was that they were not that knowledgeable about smartphones and applications. One of the participants had to be assisted by the researcher in using the application. Most of the participants also could understand what Cate was saying despite some not being very fluent in English. However, one participant had a difficult time because they were hard of hearing while another forgot to bring his reading glasses so he could not see properly. These problems however were minimized because Cate provide both text and audio versions of
the conversations.

Acceptability had the lowest score among all categories. Based on the observation checklist, they enjoyed talking to Cate but they still prefer conversations with actual humans. One participant even commented that human emotion can never be replace with robotics. Another participant said that they would only use Cate again if there was someone who could assist them, this is because the participant had an illness that limited her body movement and had bad eyesight.

For Utility, the elderly who were able to test the call and reminder functionalities were amazed that she could do that and agreed that it was very helpful to have these features.

5.2 Comparison between the Two Demographics

Figure 5 shows side by side the average scores for each category for both demographics. For the participants from the retirement home, utility and interaction had the highest scores. Based on the observation checklist, they found the call function a helpful feature. They also said that the reminder function would help them keep track of events such as taking their medicine. They mostly gave high scores for their interaction with Cate because they liked talking to her.

![Figure 5: Comparison of average scores for each category for both demographics.](image)

For the elderly living with their family, the highest score was for joy-of-use which falls under the interaction and affect category. Based on the observation checklist, they found Cate interesting and amusing. They also found it easier to use the application.

The participants from the retirement home generally gave higher scores than the participants who lived with their families. This could be due to the difference in their living conditions. Those that live with their families get to talk to them every single day while those at the retirement home often do not meet their families at all. Another reason could be due to the difference in their exposure and familiarity to technology. The elderly living with their family are more familiar with smartphones and gadgets compared to those living in the retirement home. This could also explain the large difference in scores for the embodiment category.

6 FUTURE WORK

This research is an ongoing work on developing a conversational agent that can help alleviate the social isolation and loneliness of elderly people by providing companionship to them. Based from the results, the elderly generally liked talking to Cate and liked the application. However, although they accept Cate as a companion, it is not to the point that they would use Cate every day due to certain limitations in the system and the fact that they would still prefer the companionship of another human over a machine. Nonetheless, Embodied Conversational Agent (ECA) like Cate has a potential to alleviate the social isolation and loneliness of elderly. It is important to focus on the activities (e.g. games, setting a reminder, reminding to call family and friends) first since it catches the attention of the elderly.

There are still a lot of improvements to be made as the features of Cate right now are still very basic. The dialogues must be improved since they are too static due to the limited templates used. The speech recognition and natural language processing has to be improved in order to make the conversation more natural for the elders because right now their answers are restricted to the options presented in the buttons. We also plan to add voice input and facial recognition so that we can also add other processing techniques like emotion recognition and adjust the conversations and the facial expressions of Cate accordingly. We also plan to look at other models aside from those provided by SmartBody to address the concerns regarding the robotic voice of Cate and her appearance. Since we mentioned that some of the participants also had a difficulty in understanding English, it would be also worth exploring the use of the local language instead. For the evaluation, since one of the focus is combating loneliness, use of the UCLA Loneliness Scale, the most widely used self-report measure of loneliness, will be explored.
REFERENCES


