Enjoyment, Intention to use and actual use of a conversational robot by elderly people
Heerink, M.; Kröse, B.J.A.; Wielinga, B.J.; Evers, V.

Published in:
HRI 2008: Proceedings of the third ACM/IEEE International Conference on Human-Robot Interaction

Citation for published version (APA):
ABSTRACT
In this paper we explore the concept of enjoyment as a possible factor influencing acceptance of robotic technology by elderly people. We describe an experiment with a conversational robot and elderly users (n=30) that incorporates both a test session and a long term user observation. The experiment did confirm the hypothesis that perceived enjoyment has an effect on the intention to use a robotic system. Furthermore, findings show that the general assumption in technology acceptance models that intention to use predicts actual use is also applicable to this specific technology used by elderly people.

Categories and Subject Descriptors
H.5.2. [Information Interfaces And Presentation]: User Interfaces - Evaluation/methodology.

General Terms
Measurement, Experimentation, Human Factors, Standardization, Theory, Verification.

Keywords
Human-robot interaction, technology acceptance models, eldercare, assistive technology.

1. Introduction
In the last decade, there has been an increased interest in the use of robots in eldercare [1]. The positive experiences, the many possibilities and the growing labor shortage in the industrialized world have encouraged researchers to explore this field and this particular user group. They face not only technological issues; also, the way elderly people are coping (or not coping) with new technology appears to raise challenging questions, particularly on design demands and user psychology [2-5].

In our research, we address some of those questions by exploring the factors that may influence acceptance of a conversational robot by elderly users [6]. We not only have to deal with the fact that the user characteristics of elderly people differ from the user groups that are addressed in most acceptance studies [2]; we are also facing a type of technology that brings about different aspects [7]. For example, for many users robots may not only be a piece of technology, but also more or less a personality that one might or might not accept. Besides, robotic technology could have specific aspects that other assistive technology often lacks: it might actually be fun to work or even play with.

Users might actually feel the same enjoyment they would feel when playing a game or having a pleasant conversation with a person and this might encourage them to use it.

In technology acceptance models, enjoyment is sometimes incorporated as ‘Perceived Enjoyment’, defined as ‘the extent to which the activity of using the system is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated’ [8]. Most acceptance models however, are developed within the context of utilitarian or productivity-oriented systems and Perceived Enjoyment is usually not incorporated as a major influence, while for hedonic, or pleasure-oriented systems it seems to be a crucial factor [9].

If we consider robotic systems being used as assistive technology having a place in between utilitarian and hedonic system types, the influence of Perceived Enjoyment is still unexplored.

Besides, technology acceptance models are generally based on the assumption that acceptance can be measured by mapping the influences on the intention to use a system, assuming this intention predicts the actual use of it. In our case, Perceived Enjoyment after a first experience with a system would have a predictive influence on the intention of elderly users to use it and this intention would predict the actual use of it.

The goal of this paper is to explore this concept of Perceived Enjoyment and try to establish its influence on the acceptance of robotic technology by elderly users. This acceptance is to be measured both by the intention to use the system and by actual use of it. After describing related research and theoretical concepts, we will explain how we set up an experiment in an eldercare institution to gather data on enjoyment, intention to use and actual (long term) usage of a specific type of robotic technology. After analyzing the results of this experiment, we will establish the preliminary position of Perceived Enjoyment in an acceptance methodology and set out a path for further development of an appropriate model.
2. Related work

In this section we will discuss related research on robotic technology being used in eldercare, on applying acceptance methodology to robots and on establishing the influence of Perceived Enjoyment on technology acceptance in general.

2.1 Robots in eldercare

A growing number of projects address the development of conversational robots for experiments in eldercare. A growing number of projects address the development of conversational robots for experiments in eldercare. Some studies focus on the possibilities and requirements of these ‘eldercare companion robots’ [1, 10], while other projects focus on development or on measuring the responses to it by performing experiments with specific robots.

In Japan and more recently in other countries, Wada and Shibata performed experiments with a seal shaped robot (Paro) [11-13]. These experiments showed that a robot could have the same beneficial effect on elders that a pet can have, making them feel happier and healthier. In their studies they show how different measurement methodologies can be used when studying the effect of a companion robot on elderly users.

In the US, one of the first projects with a specific robot for eldercare was done with Pearl, a robot that could actually provide some assistance to elders, although its functionalities were merely simulated [14, 15].

A more recently developed robot to be applied in eldercare is the Huggable. This robotic bear is to serve as a pet like companion, much like Paro, but with more advanced functionalities [16].

In Germany, a very sophisticated robot called Care-o-bot was and is still being developed to provide assistance in many ways, varying from being a walking aid to functioning as a butler [17]. In the most recently developed companion of this type (“Care-O-bot 3”), social and physical aid functions are represented in separate versions [18].

These different examples suggest that robots could both perform as social actors and fulfill practical functions, although the focus obviously differs within the different projects.

2.2 Measuring robot acceptance

Related research on acceptance of a conversational robot is described by De Ruyter et al [7]. It concerned a robotic interface (the iCat made by Philips), which was tested in a Wizard of Oz experiment where the robot was controlled remotely by an experimenter while it was suggested that the robot was autonomous. This experiment was done in a laboratory setting, with adult, but not elderly participants. The participants were asked to program a DVD-recorder and to participate in an online auction, by using the iCat interface. They were exposed to an introvert and an extravert version of the iCat interface to see whether this difference in interaction would lead to different scores in degree of acceptance. To measure acceptance, the UTAUT questionnaire (Unified Theory of Acceptance and the Use of Technology) was used[19]. UTAUT is a model that incorporates several influences on acceptance of technology, usually in the workplace. The aim of the study was to find out to what extent participants would use the iCat at home after having experienced it.

To see whether participants would perceive the extravert iCat to be more socially intelligent, a social behavior questionnaire (SBQ) was developed and used. The results showed that the extravert iCat was indeed perceived to be more socially intelligent and that this version also was more likely to be accepted by the user. The same robot was used in an experiment by Looije et al.[20] where it featured as a personal assistant for a small group of people with diabetes. Results showed that participants appreciated a more social intelligent agent more and had a higher intention of using it than a less social intelligent one.

It appears that research on robot and agent acceptance can be subdivided into two areas: acceptance of the robot in terms of usefulness and ease of use (functional acceptance) and acceptance of the robot as a conversational partner with which a human or pet like relationship is possible (social acceptance). The experiments with Paro were more focused on social acceptance while the experiments with Pearl and iCat (by De Ruyter et al. and Looije et al.) focused more on the acceptance of the robot regarding its functionalities. In our earlier research, in which we also used iCat [21], we concluded that both aspects need to be part of a complete robot acceptance model.

2.3 Enjoyment and intention to use

Since the first introduction of the technology acceptance model (TAM) in 1986 [22], it has become one of the most widely used theoretical models in behavioral psychology [23]. In its most basic form it states that perceived usefulness and perceived ease of use determine the behavioral intention to use a system and it assumes that this behavioral intention is predicting the actual use [19, 24-26]. The basic model has not only been used for many different types of technology, it has also been extended with other factors that supposedly either directly or indirectly influence intention to use or usage. In 2003, Venkatesh et al. published an inventory of all current models and factors and presented a new model called UTAUT in which all relevant factors would be incorporated [19].

In these models, the main instrument to measure these influences is by using questionnaires. These questionnaires consist of a number of items which can be questions or statements. Items that measure the same influence can be grouped as a measure of more general constructs. The validation of a model typically includes a long term observation of the actual use of technology, which makes it possible to relate scores on intention to use to actual usage [23].

The original TAM, related models and UTAUT were merely developed for and validated in a context of utilitarian systems in a working environment. Robotic technology used outside a working environment provides systems that might be experienced as more than this: users might have a sense of entertainment when using it. Van der Heijden [9] points out that in ‘hedonic systems’, the concept of enjoyment is a crucial determent for the intention to use it.

Of course, robotic technology in eldercare will hardly be developed just to entertain: it will be partly utilitarian, partly hedonic. But even if just partly hedonic, enjoyment could prove to be a construct that needs to be part of an acceptance model for robotic technology in eldercare.
Besides, perceived enjoyment can also be of importance in utilitarian systems, as pointed out in an extensive study by Sun and Zhang [27], although its effect on intention to use in that case could be less direct. The study mainly supports the claims by Venkatesh et al. [19] and Yi and Hwang [28], that Perceived Enjoyment has no direct influence on Intention to use, but that it can influence Ease of use and Usefulness. Still the study does also recognize that this is not a general claim for all types of systems. Indeed this could work very differently for robotic systems used by elderly people.

An acceptance study also including perceived enjoyment by Chesney, concerned the use of Lego Mindstorms development environment by Mindstorms hobbyist [29]. The study, based on the viewpoint that this concerns a partly hedonic, partly utilitarian type of system, confirms perceived enjoyment having just an indirect effect on intention to use.

We may conclude that literature on acceptance models in general does attribute some influence to perceived enjoyment in systems that are partly or totally hedonic. Since socially interactive robots may be experienced as hedonic systems, this means perceived enjoyment could be of some influence. When we consider social acceptance also to be a factor, especially with conversational robots, this means robotic systems differ from the systems described in acceptance model literature so far and the strength of the influence of perceived enjoyment is still very much uncertain, especially in the context of eldercare.

Focusing on social acceptance, relating the concept of perceived enjoyment to the intention to use a conversational robot brings us to another acceptance model issue that is of interest to us. An important aspect of these models is the assumption that behavioral intention to use determines actual use, is correct. This means that if we want to make assumptions on influences on intention to use, we have to develop a way to measure perceived enjoyment, intention to use a system and actual usage. In the next section we will present an experiment that has been set up for this purpose.

3. Experiment

With regard to social acceptance we researching the influence of perceived enjoyment on acceptance. This means in terms of acceptance methodology we want to establish its influence on the intention of users to use the system. In our case this means we want to see if elderly people after a first impression are more willing to use a socially interactive robotic system. Besides, we want to find out if this intention really predicts the amount of usage of this type of system by this type of user.

Thus, there are two hypotheses we want to test in an experiment:

H1. The more people perceive a robotic system to be enjoyable, the more they intend to use it.

H2. The more people indicate they intend to use a robotic system, the more they will actually use it.

We set up an experiment with the intention to establish (1) the relationship between Perceived Enjoyment and Intention to Use and (2) the relationship between Intention to Use and Usage.

Measuring perceived enjoyment and intention to use demands a setup in which there is a small test in which people get a first impression, while measuring usage demands a setup in which people can be observed using or not using the system over a certain period.

3.1 Robotic system

The robotic agent we used in our experiment is the iCat ("interactive cat"), developed by Philips, also used in the experiments by De Ruyter et al.[7] and Looije et al.[20] and within our own project [21]. The iCat is a research platform for studying social robotic user-interfaces. It is a 38 cm tall immobile robot with movable lips, eyes, eyelids and eyebrows to display different facial expressions to simulate emotional behavior.

There is a camera installed in the iCat’s nose which can be used for different computer vision capabilities, such as recognizing objects and faces. The iCat’s base contains two microphones to record the sounds it hears and a loudspeaker is built in for sound and speech output.

For this experiment, we used a setup in which the robot was connected to a touch screen as is shown in Figure 1. In earlier setups we had used the robot in a setup in which it was voice controlled, but for smooth interaction this demanded a hidden operator to control the robot (a so called Wizard of Oz setting). This would not allow us to leave the robot for public use for a few days, so we chose the touch screen control.

Figure 2. Test participant using iCat with touch screen

It could be used for information and for fun: the participants could ask for weather forecast, a television program overview or a joke by pressing the appropriate choices from a menu on the screen. The information was then given with pre-recorded speech by the iCat, for which we used a female voice. The recording was done with a text to speech engine.
3.2 Method
We designed an experiment in two eldercare institutions in the city of Almere in the Netherlands with the first part consisting of a short test, during which participants were to meet a robot and work with it for a few minutes individually.

3.2.1 Subjects
Participants were recruited both by eldercare personnel and by students. Their age ranged from 65 to 94, while 22 of them were female and 8 were male. Some of them lived inside the eldercare institutions, some lived independently in apartments next to the institutions.

3.2.2 Procedure
Participants were brought into a room were they were alone with the iCat and one researcher. They did not get any specific task, but were instructed to simply play with the robot for about three minutes. After the participants finished this session they were brought to another room where they were given the list with statements to which they could reply. They could ask for help if they were unable to read the statements.

After these sessions were completed, we left the robot for public use in a tea room near the entrance: a place where most of the population of the eldercare institution would see it. On the screen were buttons with the names of the ten test session participants and one extra button saying “I’m not listed”. Passers by were informed by a note that anyone could use the robot and that they could start a session with pressing the button with their name on it. If their name was not listed, they could use the “I’m not listed” button.

3.2.3 Instruments
We used a questionnaire that consisted of a list of statements that participants could reply to in a five point Likert scale (totally disagree – disagree – don’t know – agree – totally agree). Table 1 shows the used statements on the constructs Intention to use (ITU) and Perceived enjoyment (PENJ).

### Table 1 – Statements (translated from Dutch) for Perceived Enjoyment and Intention to Use

<table>
<thead>
<tr>
<th>Construct</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITU</td>
<td>I think I’ll use iCat the next few days</td>
</tr>
<tr>
<td>ITU</td>
<td>I am certain to use iCat the next few days</td>
</tr>
<tr>
<td>ITU</td>
<td>I’m planning to use iCat the next few days</td>
</tr>
<tr>
<td>ITU</td>
<td>I think I’ll use iCat for this amount of minutes: 0</td>
</tr>
<tr>
<td>PENJ</td>
<td>I enjoy iCat talking to me</td>
</tr>
<tr>
<td>PENJ</td>
<td>I enjoy doing things with iCat</td>
</tr>
<tr>
<td>PENJ</td>
<td>I find iCat enjoyable</td>
</tr>
<tr>
<td>PENJ</td>
<td>I find iCat fascinating</td>
</tr>
<tr>
<td>PENJ</td>
<td>I find iCat boring</td>
</tr>
</tbody>
</table>

During the days the iCat was available for use to anyone passing by, the system made video recordings as soon as it was used through the camera in its nose. Furthermore, it kept a log of the start and end times of individual user sessions. The end time was either the time a user actively ended his session or if it was not used for 90 seconds.

By comparing the video footage to the log, we could later check if users had pressed the right button.

4. Results
The test session and the questionnaire were completed by 30 participants. In analyzing the replies to these statements, we used Cronbach’s alpha to test the reliability of the constructs. In psychology, an alpha of 0.7 and higher is considered acceptable [30]. As table 2 shows, the constructs were highly reliable.

### Tabel 2 – Cronbach’s alpha

<table>
<thead>
<tr>
<th>Construct</th>
<th>Cronbach’s alpha</th>
<th>N of items</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITU</td>
<td>.947</td>
<td>4</td>
</tr>
<tr>
<td>PENJ</td>
<td>.836</td>
<td>5</td>
</tr>
</tbody>
</table>

Regarding the long term part of the experiment, we analyzed the video footage and the log, and compared these to find out if users pressed the correct button. We found that there were 122 full sessions of which 70 were from test session participants. Users that did not belong to this group did not always use the “I’m not listed” button: 17 of 52 of their sessions were started by them using the name of one of the test participants. We omitted the usage data of these users from the usage scores. The test session participants however, always started their session with their own name.

Table 3 shows the mean scores on the statements (scores of 1 to five were attributed to the replies) and on usage, measure both in the number of sessions (‘Times’) and the total amount of minutes for each participant.

### Table 3 – Descriptive Statistics for on Perceived Enjoyment and Intention to Use (n=30)

<table>
<thead>
<tr>
<th></th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Times</td>
<td>0</td>
<td>9</td>
<td>1.90</td>
<td>2.339</td>
</tr>
<tr>
<td>Min.</td>
<td>0</td>
<td>16</td>
<td>4.90</td>
<td>4.634</td>
</tr>
<tr>
<td>ITU</td>
<td>1,00</td>
<td>4,75</td>
<td>3,2250</td>
<td>1,37143</td>
</tr>
<tr>
<td>PENJ</td>
<td>1,60</td>
<td>5,00</td>
<td>3,6667</td>
<td>.84418</td>
</tr>
</tbody>
</table>

In fact 23 of the 30 test session participants did use the system later, varying from one to nine times. It also shows that the session length varied from one to five minutes and that the maximum total amount of minutes per user was 16 and the maximum times of usage was 9. The test sessions were not included in these counts.

The correlation between the constructs is presented in table 4. It shows significant correlation scores between Perceived Enjoyment and both Intention to Use and usage in minutes. Also it shows a significant correlation between Intention to Use and times of use, but the score is exceptionally high between Intention to use an usage counted in minutes per participant.
Table 5 shows the results of our analysis, with the Independent variable predicting Intention to Use and Intention to Use predicting the dependent variable in the first column. The scores show that a linear regression analysis was used when modeling several constructs [19, 23], but since we have a small number of participants and just one predicting factor per hypothesis, we performed a simple linear regression analysis.

Table 5 – Linear regression: t-scores for Perceived Enjoyment predicting Intention to Use and Usage in Times and Minutes

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Dependent variable</th>
<th>Beta</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PENJ</td>
<td>ITU</td>
<td>.420</td>
<td>2.449</td>
<td>.021(*)</td>
</tr>
<tr>
<td>ITU</td>
<td>Times</td>
<td>.413</td>
<td>2.400</td>
<td>.023(*)</td>
</tr>
<tr>
<td>ITU</td>
<td>Minutes</td>
<td>.625</td>
<td>4.236</td>
<td>.000(**)</td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

In acceptance modeling, usually a multiple regression analysis and path analysis is used when modeling several constructs [19, 23], but since we have a small number of participants and just one predicting factor per hypothesis, we performed a simple linear regression analysis.

Table 4 - Pearson correlations for Perceived Enjoyment, Intention to Use and Usage in Times and Minutes

<table>
<thead>
<tr>
<th></th>
<th>ITU</th>
<th>PENJ</th>
<th>Times</th>
<th>Minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITU</td>
<td>Pearson Correlation</td>
<td>.625(**)</td>
<td>.420(*)</td>
<td>.413(*)</td>
</tr>
<tr>
<td>SIG. (2-tailed)</td>
<td>.000</td>
<td>.021</td>
<td>.023</td>
<td>.000</td>
</tr>
<tr>
<td>PENJ</td>
<td>Pearson Correlation</td>
<td>.363(*)</td>
<td>.420(*)</td>
<td>1</td>
</tr>
<tr>
<td>SIG. (2-tailed)</td>
<td>.049</td>
<td>.023</td>
<td>.120</td>
<td>.000</td>
</tr>
<tr>
<td>Times</td>
<td>Pearson Correlation</td>
<td>.861(**)</td>
<td>.413(*)</td>
<td>.290</td>
</tr>
<tr>
<td>SIG. (2-tailed)</td>
<td>.049</td>
<td>.023</td>
<td>.120</td>
<td>.000</td>
</tr>
<tr>
<td>Minutes</td>
<td>Pearson Correlation</td>
<td>.861(**)</td>
<td>.413(*)</td>
<td>.290</td>
</tr>
<tr>
<td>SIG. (2-tailed)</td>
<td>.049</td>
<td>.023</td>
<td>.120</td>
<td>.000</td>
</tr>
</tbody>
</table>

For example, the fact that elderly people tend to respond to a robot as if it were a human being [31] opens the possibility of the sense of social presence being of influence.

Regarding Intention to Use and Usage, the acceptance model assumption of the first being predictive towards the second has been proved to be applicable to robotic systems used by elderly people. This encourages the quest for a new model that can be used in this specific context, with this specific technology.

Future research needs to be careful in applying these findings, the robot in our experiment has only been available for five days and only ten participants were fully involved. The development of a robot acceptance model including all relevant factors, demands usage behavior observations, involving much more participants. But despite this need for larger numbers, we may conclude that this type of experiment delivers useful data to validate the model. It delivers objective data on usage, where in the validation process of models often usage is measured by less reliable self reported usage.

Something that might be taken into account in further research might be a phenomenon observed while analyzing the video footage and identifying the users. Of the 52 sessions by users who did not participate in the test sessions, there were 27 sessions by 15 users who earlier indicated that they were not interested in participating in the test sessions. The images revealed every one of them was sitting, playing with the robot at a moment that the tea room was empty. This means that in our case, not wanting to participate in a test session does not imply not being interested in (or at least curious about) the system.

It could very well be that the possibility of feeling embarrassed and not a lack of interest prevented them from taking part in our experiment. It would be interesting to see if some of those people would become regular users and we may indeed have more valuable data if we somehow could tempt participants like these to take a closer look at the robot – perhaps not in the present of researchers - and fill out the questionnaire. Perhaps in further research the challenge to fit those people in our methodology somehow has to be met. Besides, we may need to be more inventive in finding ways to avoid the impression that test participants may have to face some kind of embarrassment.

Regarding the implications on the design of interactive robots, this study shows the importance of non-functional aspects that may raise the level of enjoyment for elderly participants. Further research might focus on the design aspects that are increasing perceived enjoyment both in general and more specifically for elderly users. Besides, it also may address the way a robotic companion is presented to future elderly users. It could very well be essential to communicate that a social robot is not just an assistive device, but also a very enjoyable companion.

Considering the amount of data gathered with this type of experiment, we conclude that it was an accurate method: despite the limited functionalities, we collected a useful amount of usage data in just a few days.

5. Discussion and conclusions

Both hypotheses could be confirmed on the basis of a simple linear regression analysis. The high correlation and regression scores between Perceived Enjoyment and Intention to Use indicate that this influence is strong. This means perceived enjoyment needs to be part of an acceptance model applied to robotic systems to be used by elderly people.

To construct such a model, of course the relationship of perceived enjoyment with other influences like perceived usefulness and perceived ease of use should be established. In fact, there are many regular model constructs (Attitude, for example) that it could be related to. Future work needs to address these relationships, but it also needs to map possible factors that are not considered major influences for utilitarian systems being crucial in this context.

6. Acknowledgments

This work was supported in part by the Hogeschool van Amsterdam and in part by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020 (Cogniron).
We like to express our gratitude to the staff and caretakers as well as the test participants of “De Kiekendief” and “De Overloop” in Almere for their cooperation. Also we like to thank Willem Notten, Bas Terwijn and Rogier Voors for their work on programming the system and Rick van Middel, Albert van Breemen and Martin Saerbeck for their support.

References

