A Personal Approach: The Persona Technique in a Companion's Design Lifecycle

Joana Campos, Ana Paiva

A personal approach: the *persona* technique in a companion's design lifecycle

Joana Campos and Ana Paiva

Instituto Superior Tecnico - UTL and INESC-ID, Portugal
joana.campos@ist.utl.pt
ana.paiva@inesc-id.pt

Abstract. *Artificial companions* are a new type of technology that is changing the way people interact with intelligent systems, by trying to build and sustain long term relationships with users. To evaluate such systems the typical usability methods for system evaluation are not enough, due to the need of accessing aspects such as social behaviour, emotional sensitivity and personalized interaction over long periods of time, with very specific users. In this paper, we describe the (full) design cycle of a companion system, arguing that a user-centred approach is the more appropriate, taking into account the referred specific issues. As such, to help us focus on developing a companion system for the prospective users, we based our design in the archetype of actual users – *a persona*. Furthermore, we brought this same concept into the evaluation phase in order to access the companion performance in a long-term interaction.

1   Introduction

During the past few years we have witnessed a new boost in the area of intelligent agents applied to human-computer interaction (HCI). A new kind of technology is gaining ground as it tries to establish "social relationships" with users, rather than simply interacting with them [2]. Embodied as robots, 3D synthetic characters, or even interactive toys, an *artificial companion* should be able not only to promote more pleasing and natural relationships with users, but also to build and sustain such relations over long periods of time.

Companion systems are designed to assist or help users in a specific task and at the same time to act socially in an appropriate fashion [7]. Furthermore, this technology attempts to achieve long-term robustness, and as a result it should encompass some capabilities, such as affect sensitivity, memory or learning [7]. The Senior Companion system [8], for example, helps elders to organize their life memories, while it provides comments and tries to establish an emotional connection. Another suitable example is the Health and Fitness Companion [20], which tries to persuade the user to engage in

a different and healthier lifestyle. It does so by providing advice and guidance throughout the day about combinations of diet and exercise.

Indeed, *companion systems* are a new type of technology that allows one to manage his/her life and personalize the interaction accordingly. Furthermore, these systems are task specific and user focused, whereupon to address design issues it is crucial to adopt a methodology that iteratively involves the user in its design process. Therefore, we argue that user-centred design methods convey this process in a dynamic way, wherein the user participation is essential throughout the whole process both from the requirements analysis to the evaluation phases. However, when users are not present, there are still a lot of decisions to be made that should follow the same paradigm.

To overcome this issue and to help us focus on developing a companion system for the prospective users, we based the design of our *companion* on an archetype of actual cohorts – a *persona*. This technique relies on a character that is built upon information from potential users who participated in early stages of design. The underlying idea of *personas* is to increase user focus and awareness, as they are specific assumptions about a set of people and also to make the decision criteria explicit [18][24].

In this paper we describe the design cycle of a *companion* system (fig.1) focusing on the essential processes of user-centred design (UCD): plan and identify context of use, gather requirements, prototype design and evaluation [22]. We would also like to emphasise that the use of UCD enables a shorter development lifecycle by guiding the implementation process.

However, when evaluating *companion* systems there are two key issues to consider [2]. First, relationships are long-term built over many interactions and secondly relationships are persistent, emotional and personalized. Undoubtedly, the memory is an essential component for the agent to show its capabilities according to the two previous issues. As such, it should be primarily populated with a considerable number of interactions. Yet, this issue would turn the user-centred design for *companions* a very expensive and extremely long process. That is related to the fact that long-term evaluations between participants and companions are costly and hard to manage.

As [3] suggested companions technologies require new models for evaluation. For that reason, we suggest to use the concept of *persona* on the early evaluation stages. As the *persona* is a complete and carefully defined character it would be possible to fill the *companion's* memory with the tailored *personal* experiences. We believe that by following this paradigm we can simulate previous interactions and access the agent's performance, whereas we are simulating a long-term interaction in a short duration evaluation.

2   Companion Context of Use and Design Overview

Throughout this paper we will describe the development of a companion system – MAY (my Memories Are Yours) – created to assist a teenager user on self-reflection and daily companionship about what happens in his/her life [5]. This system tries to preserve the user's expressivity in analogy with a traditional diary.
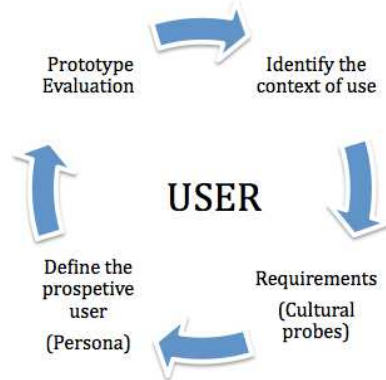
Fig. 1. The companion design cycle

To try to achieve that goal it is necessary to take into account that language is the currency of most human social processes [10], so the interaction is based on that issue. Moreover, the main component of this *companion* is its memory, as we believe to be not only an essential aspect to improve intelligent systems in general, but also an important feature to help our *companion* behave more naturally in a social setting.

When planning the design process, we used a set of (what we believe to be) suitable methods for guiding a *companion* development focusing on a UCD perspective. First, we used the cultural probe method for gathering information about user's tasks and to analyse their needs in a less obtrusive way. From that data we were able to identify and formally define a *persona*, which guided the subsequent stages of the *companion* development. With the gathered information, we were able to build a companion prototype, which was also evaluated. This time, we used the *persona* as a central element of the process. The role of the potential users that participated in this phase was to classify whether the memory allowed the agent to produce appropriate responses according to its content (persona's experiences), or not.

3   Personas

Persona is an archetype of actual users with well-defined characteristics and it was first introduced in HCI community by Cooper [11]. Their use helps designers focus on prospective users of a system or interface, taking into account their needs and goals [9]. These imaginary characters must be defined according to data retrieved from initial investigation of the target user.

Personas have various attributes such as name, age, gender, educational achievements, profession, possessions, likenesses, family, occupation, friends, pets, tasks, life stories, goals, among others [19]. They are "imaginary" users that designers consider while developing a product. This technique arises from the difficulty of

continuously work closely with potential users on a possible new product, especially in large organizations [19].

The use of personas has recently become popular worldwide as a design method. For instance, it was used for the development of MSN Explorer products [24]. This success can be attributed to its potential of engaging designers during the prototyping phase. Researchers argue that people more easily remember stories and narratives than technical reports. Consequently, *personas* with their stories can persist on the designers mind set. Furthermore, personas act as a mechanism to enhance attention and help on the organization of data [19] carried out during the design phase.

This idea is also prevalent even in the design of TV series. It is frequent to see people believing that TV characters are real and see their life as realistic. Additionally, people are keen in discussing details and make predictions about characters behaviour. These predictions are made according to a set of characteristics, which they have gathered from seen episodes. On the other side, the actor builds his character and exercises a story where the character may fit. Certainly, the actor adds some details to help him behave more naturally [19], as well as, the designers make assumptions about the created persona.

Some studies [9][24] evaluated if the presence of the *persona* factor had a positive effect on the final design comparatively to the absence of it. Their studies underlined that the use of personas helped designers on picturing the user and get to a consensus. However, *personas* should be used for a particular effort and should not substitute other techniques for user-centred design. Instead, they should complement them by enhancing user focus. On the whole, this is a good approach for complementing and guiding not only the design process, but the test phase as well [19].

If the persona is efficient in engaging designers to focus on the users, it should be equally effective on engaging users, when they perform tests with some system or interface. That issue was already raised [18] as the author used personas to test the usability of a web prototype.

4   Gathering Information about the Users

When starting the design process, our initial purpose was to uncover the practices through which people experience memory sharing and discover how they perform activities that the system tries to assist. But how could we analyse 'real life' without losing personal expressivity?

To achieve a 'close view' of daily life without blurring the essential aspects (the one's that we had not thought of) we moved away from the controlled environment found in a lab, sending a probe to potential users and waiting for its return.

Cultural probes are an interesting collecting process, wherein people agree to share personal aspects of their lives. The probe exercise gives necessary detachment from the designer, and provides information about how people use the given materials to express themselves in an unobtrusive way [13]. Probes do not directly lead to designs, as we could experience, but make us aware of details, which are valuable in inspiring design ideas [16]. We used this approach as a starting point to satisfy our intentions.

Our experiment was focused on teenagers, mainly ones that have just started their academic life in the university or were about to start it. We wanted to disclose what they are willing to tell to a companion, which things are important in their life and which are of their concern. How 'stories' of their life are told, and consequently organized in words, and which are their routines. Further, we wanted to analyse how people used the provided materials.

4.1   The Probes

We gave to the participants an envelope (Fig. 2) containing ten tasks to perform, a blue pen, coloured pens, post-its, stickers with emotions and a notebook – their *companion*.

Each task had clear instructions of what was needed to do, yet not restrictive about how to perform them. The first task aimed at creating a first bond with the companion requiring an image of it/she/he. The following tasks, tried to 'extract' ways of interaction and sharing experiences (memories) process. People had liberty to deal with the requests as they preferred.

Fig. 2. The probe pack

The 14 participants (seven boys and seven girls) from ages 16 to 18 were first-year students of different universities and high school finalists. They took the probes and returned them 3 weeks later.

Probes should not be used to produce scientific analysable data [16], therefore, it is common to give participants ambiguous and absurd tasks to create detachment from preconceived definitions of the world. Although our designed tasks (for more information please see [6]) were relatively straightforward, they called to a subjective interpretation from us, and a deeper thought from the participants. However, we could experience that the most structured questions/tasks were not the ones that produced more accurate responses and that give us a clue on how to analyse the probe material. Tasks that ask them to describe episodes of their lives had a lot of (more) valuable information.

The most relevant data extracted concerns the companion image, what users consider important in their lives, as well as, the process of remembering described in a diary form.

4.2   What Was Probed?

This experiment provided an engagement with the participants, allowing us a 'close view' of their lives. The 'storytelling tone' was constant in their responses, as they literally wrote a diary. However, a direct interaction with a companion was not left out, and during the experiment some subjects really 'talked' to MAY. In the light of this, we established that the communication between user and companion would be by text to preserve the expressivity we found in the data. Then, the central point of the interaction would be dialogue, which facilitates the sharing of knowledge and individual's information.

Moreover, the probes enhanced a conversational behaviour as suggested by [14]. There is a focus on talking about personal experiences and relationships, as well as the emotional responses involved in those interactions. The personal experiences fit in one of the four important dimensions of younger's life: love, sport, leisure and school. The way in which the given materials were used did not surprise us, but stood out the necessity of a feature to enable attaching objects, like images or sounds, to the written text. Still, the probes oriented the design of the interface and a first prototype was modelled.

5   Defining the Prospective Users

The collected answers from the cultural probes, revealed to be an insight of each participant and its combination brought up a "conceptual model" of the prospective group of users and how the system should behave.

Following the *persona* design technique, we created a fictional user, named Amy (Table 1 and 2). The data extracted from one probe was our starting point to create that fictional person, which behaviour was interactively built based on real data from the rest of the potential users, who have participated on the probe study. The created character aggregated a set of characteristics from several people (participants in the probe study) and not from just one. Our aim at creating Amy was to guide the companion's design both its interface and implementation. Further, we would like to extend the use of personas to help us surpass the long-term issue, based on the engagement that this tool can provide. More details on this matter are explained on section 7.

The persona Amy is 18 years old and is on her first year of Computer Science in the university. She likes summer time and being with her friends whenever is possible. April is her best friend, and they usually go to the cinema together and out for dinner once in a while. As many teenagers one of her goals is passing the driving test and of course be successful in her studies at the university.

Table 1. Persona details and life stories. The enumerated characteristics were suggested by Grudin [19] and completed with the probes' information.

Persona Definition		
Ethnography	Name	AMY
	Age	18
	Gender	Female
	Socioeconomic Status	Middle Class
	University	Blind review
Details	Family	Amy lives with her parents and sister. She always have fun with them and they are always there to back her up
	Friends	Friends are important in Amy's life, and them are in her list of concerns. Her best friend is April, they used to go to the cinema, to go shopping or have dinner/lunch once in a while. They try see each other every Tuesday or Friday.
	Pets	She has a dog to take care.
	Sport	Every time she plays volleyball, she plays with her heart. She totally loves it and for her it's a great escape from school.
	Likes	Amy is a summer person. She also likes listening music, travelling, camping, school (why not), eat good food, going out at night with her friends. She prefers places like Hard Rock Café, but Bairro Alto is one of her choices too.
	Dislikes	Wake up early, public transports (school-home), domestic tasks at home.
	Goals	Pass the driving test, get good grades at university subjects, to be in good shape, wake up on time for school
	Tasks	Take care of the dog; at Tuesdays and Fridays is her turn to clean up the kitchen; clean up her room weekly; go to volleyball practices; go to drive lessons

6   The Scenario

As it was described before, the companion MAY, is an agent created to assist a teenager user on self-reflection about what happens in his/her life. The interaction

Table 2. Persona Life Stories. (continuation of table 1)

	Persona Definition
Life Stories	"In Summer 2008 I went on a camping trip to Açores with some people from school. Against all odds it rained a lot! The worse came later… At the middle of the night some tents collapsed and we have to go our 'neighbors' tens to spend the night. In the morning all campsite had been flooded and our clothes and stuff were soaking."
	"Today is my 16 birthday and I got my first car – a Microcar. I am so happy and excited. It is like a dream come true. I cannot wait to show it to all my friends."
	"Today was a normal day at school. Slightly boring… don't really want to work. But I went to one beach for lunch with friends, really good moreover."

between them is through dialogue, in chat like interface, which allows the agent to collect the user's experiences and save them in a diary form (or a timeline). The memories (or events) stored in the companion's memory constitute a kind of "shared memory" between the two [5], which can be compared to an affective diary [25].

To produce an adequate response, the agent starts by searching its memory for anything appropriate to say. It looks for active goals, past events with some relevant information for the current situation or even to go beyond the present and infer future plans. This allows increasing the agent's responsiveness, that is, listeners make empathetic and contextually grounded responses to what the speaker is saying. This process is considered not only to enhance believability of conversations, but also serve engagement and intimacy in relations [4] making it an enjoyable two-way interaction.

7   Persona and Companion Early Evaluation

As noted earlier, new systems that try to change how people interact with technology making them develop behaviours towards relationship building [2], require new methods for evaluation. In the context of the COMPANIONS[1] project, for example, researchers outlined a new evaluation paradigm to measure their companions' conversation quality [3]. They suggest the appropriateness annotation as a measure to evaluate the companions' conversational performance, along with objective and subjective measures. Others [26] have explored the affect and emotions in interactions as part of user experience as the interest for emotional design [27] has gained interest.

In our work, we are particularly interested if people can identify cues in dialogue that may elicit some kind of behaviour that could demonstrate social capabilities in the agent's discourse. Furthermore, we wanted to see if users perceive the agent as intentional, based on the result of previous interactions stored in a timeline. As such, a User-Centred method [3] seemed appropriate for our study. This allowed us to acquire

---

[1] http://www.companions-project.org

subjective impressions and opinions by the participants collected in Likert based surveys [3].

For a companion system showing user acquaintance during a dialogue interaction it is necessary many interactions to capture and store information. However, in early stages of the design, such kind of evaluation is costly and hard to manage.

So, we brought the concept of *persona* into the evaluation stage in order to assess the companion's memory performance in a long-term interaction, but in a short-time period. As such, we created a *persona*, called Amy (already described in section 5).

As an attempt to simulate the long-term issue, we filled the agent's memory with enough information to perform in a real situation. We introduced memories of experiences into the system, taking into account the life story we had created for Amy, supported by the idea that models of fictional persons can be engaging as real people [18]. This process was done with a help of an authoring tool to define a persona according the specified parameters, which was built by us. Using this tool, it was possible to populate large "slots" of the memory, according to activities or events that we thought to be acceptable with respect to the persona's definition.

In the companion evaluation described in this section, we tried to verify if the presence of previously shared memories between the user and the agent enabled the development of specific kind of relationship between the two, in comparison with the absence of such element.

Thus, we evaluated if *an agent that is capable of indexing user's experiences and using that temporally grounded information in a social interaction, will make users to establish a stronger companionship relation with the agent based on what the agent knows.*

7.1   Evaluation Procedure

A total of 90 participants (72 male, 18 female, aged 18-24) took part in the experiment. All of them were undergraduate and the study was available through an online questionnaire. The questionnaire conducted the participants through three stages (see fig. 3). In the first two stages the subjects had to witness an interaction between Amy and MAY, which they should observe carefully. After each interaction they would be asked about MAY's behaviour. The last stage aimed at measuring how MAY's attitudes induced the relationship between Amy and MAY. For this last stage we applied an established friendship questionnaire. In each stage, when inquired, the subjects had to classify statements in agreement with a five-point Likert scale (1 - I strongly disagree; 3 - I don't agree or disagree; 5 - I totally agree), coupled with a justification to validate the response.

Notice that none of the subjects knew what the system did, neither if it should act as a companion nor an agent that saves personal facts about the user through shared memories. Plus, the interactions between Amy and MAY reflected some of the agent's capabilities of using information collected during previous interactions with the system. The full questionnaire and interactions between Amy and MAY are presented in [6].
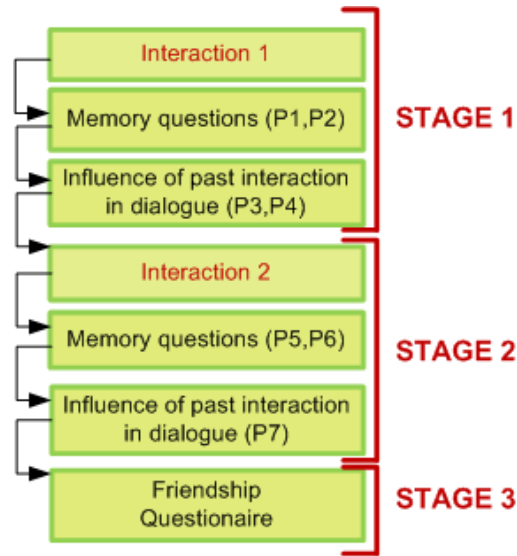
Fig. 3. Evaluation Flow

7.2   Manipulation

We conducted a between-groups experiment to evaluate the raised hypothesis. Thus, one group of participants was exposed to the memory condition and the other group to the no memory condition. The online questionnaire automatically randomized the participants to one of the conditions, which express the following situations:

- *memory* – there is no modification in MAY's architecture. There is a link to the Agent's Cognitive Features Module. This module is responsible for bringing back relevant information for the current situation.
- *no memory* – MAY's architecture was manipulated and there is no access to the Agent's Cognitive Features Module.
  - However, in this condition, MAY still knows the user's name and maintains the pro-active behaviour based on the design of the dialogue system. That is, it still 'cares' about the user by asking intentionally generic questions to get information about user's day.

7.4   Results

The collected data was analysed using the Mann-Whitney test to compare the differences between the two conditions.

For clarification, on the very right-hand column of each table (from stage 1 onwards) we provided r as a measure of the effect size, i.e. how significantly different the two conditions are. Although r may be interpreted as a correlation, in this study we use it to measure the strength of the experimental effect [25], which in this case was the presence of memory in the companion's responsiveness in each of the variables (questions 1 to 7).

The test-statistics (U) was converted into a z-score by SPSS and the r was calculated using the following formula: $r = z\text{-score}/\text{square}(N)$, wherein N is the size of the sample. Consequently if z-score has a negative value the r will have a negative value as well. One may argue that means that the experimental group was worse off on these areas. But it does not, as it is not considered as a correlation factor. The issue is that variables were inversely encoded in SPSS. However, it is irrelevant in this case, as we want the absolute value of r.

Stage1. In the first stage of the experiment the users watched an interaction between Amy and MAY, where the agent shows that he knows some predefined goals in Amy's life. Below are some examples of the type of interaction (memory condition)[2]:

```
11  MAY >> How was the Cranberries concert last night?
    Amy >> It was awesome! She sang all the songs!

...

25  MAY >> I'll be waiting.

        Did you go out for dinner two days ago?

27  Amy >> Yes. It was my sister's birthday. ...
```

Rather, in no memory condition such acquaintance is not verified, but the agent displays an empathetic behaviour embedded in the dialogue system:

```
11  MAY >> How was your day yesterday?

    Amy >> It was awesome! I went to the Cranberries'
    concert!

...

25  MAY >> I'll be waiting.

        Tell me your last and important memories.

27  Amy >> Saturday I went out for dinner. It was my
    sister's birthday. (...)
```

---

[2] The full interaction is depicted in [6]

Figure 4 depicts a summary of the results obtained for the different questions at this stage of evaluation and exhibits the clear differences between the two conditions (see also Fig.6 for a graphical visualization of the descriptive statistics)

As the results show, users recognised the presence of knowledge about goals (Q1, Q2) and the results between conditions were significant (p < 0.001). These results may seem obvious, but we would like to emphasise that the users didn't know anything abut the system nor how it should behave. Nevertheless, the most relevant data refers to Q4, where the user recognised the information depicted in the dialogue as temporally grounded.

However, no memory condition had a wide discrepancy of opinions that ranged within the five points (Mdn[Quartiles]=3.5[1,5]), when the participants were asked about the influence of the past (Q4). This dispersion of values is due to the agent knowing the user's name. For some participants that fact was a sufficient (in the absence of other) for the existence of prior interactions. As one user justified "MAY knew to whom was speaking, as she referred as Amy".

Cognitive Features	Id	Question Statement	Descriptive Statistics memory (N=46)	no memory(N=44)	Mann-Whitney differences between conditions
			Mdn[Quartiles]	Mdn[Quartiles]	
Tracking Goals	Q1	MAY knew about the event "go to the cranberries' concert"	5[4,5]	2[1,3]	U = 129.000 p < 0.001 r = -0.777
	Q2	MAY knew about the event "go out have dinner"	4.5[4,5]	2[1,3]	U = 232.500 p < 0.001 r = -0.68
Influence of past Interactions	Q3	MAY knew about the event "go out have dinner" because Amy had told it in any past interaction	5[4,5]	1[1,3]	U = 245.500 p < 0.001 r = -0.679
	Q4	Past interactions have influenced the current one	5[4,5]	3.5[1,5]	U = 527.500 p < 0.001 r = -0.452

Fig. 4. Mann-Whitney statistics for the first set of questions

Stage 2. In the interaction portrayed in the second phase of the evaluation, the memory condition enhanced the agent's acquaintance about some details of common or regular events in Amy's life.

```
12   Amy >> Yes. In a while I'm going to walk Teddy.
13   MAY >> Are you going to the park?
...
16   Amy >> This evening I'm going to the cinema
17   MAY >> Are you going with April, Amy? ...
```

Oppositely, in the "no memory" condition the agent keeps answering to Amy with general sentences that could fit in most situations, using the pattern-matching mechanism offered by the dialogue system.

```
12  Amy >> Yes. In a while I'm going to walk Teddy.
13  MAY >> Sounds good to me.
...
16  Amy >> This evening I'm going to the cinema
17  MAY >> I would like to go too ...
```

Cognitive Features	Id	Question Statement	Descriptive Statistics		Mann-Whitney
			memory (N=46)	no memory(N=44)	differences between conditions
Virtual Sensing	Q5	MAY knew something about the action "walk Teddy"	4[3,5]	3[1,4]	U = 589,000 $p < 0.001$ r = -0.370
	Q6	MAY knew something about the action "go to the cinema"	4[3,5]	4[2,4]	U = 722.500 $p < 0,05$ r = -0.256
Influence of past Interactions	Q7	Past interactions have influenced the current one	5[4,5]	4[4,5]	U = 797,500 $p < 0.051$ r = -0.206

Fig. 5. Mann-Whitney statistics for the second set of questions
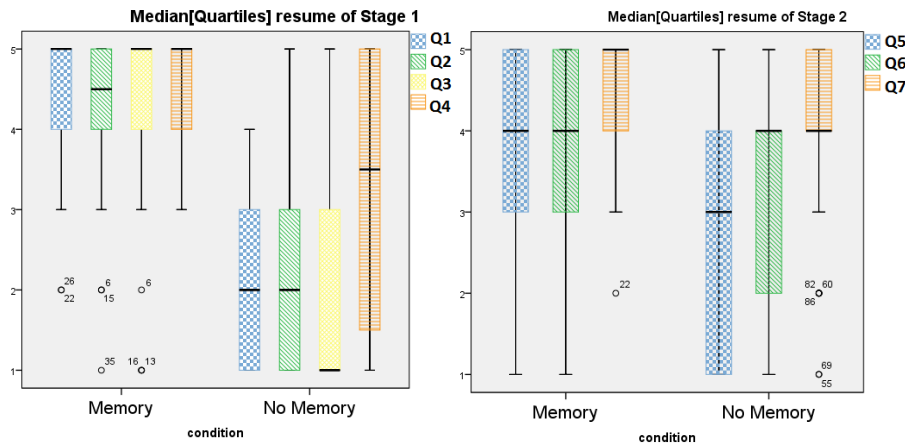

Fig. 6. Descriptive statistics for Stage 1 and 2

Once again the results in this stage may seem obvious, but users did not know anything about the system or how it should behave.

In both stages (1 and 2), specific elements in the dialogue from the stored information were recognized as intentional and as a result of past interactions. Yet, at stage 2, such statement is not so obvious as the medians of both conditions are equal or very close to each other. The explanation for significant results in this stage is based on the theory behind Mann-Whitney test. Mann-Whitney is not a median test, but a rank test. That is, it calculates the difference between two distributions of data. There is a difference between the distributions, which is supported by the quartiles (as depicted in Fig. 5).

Nonetheless, the *r* describes whether such significance is observed by chance or not. It tells us that our experimental condition had a small to medium effect size. These results may have happened for two reasons: 1) the agent's replies may have lead to a misunderstanding by the users. The responses make them feel that the agent 'knew' something about some characteristics of the activity; 2) by watching the previous interaction at stage one.

The participants probably had expectations about the second interaction, which made them think that the agent took conclusions similar to their own assumptions. According to user's justifications "MAY recognizes cinema as a fun activity because he replied that he would like to go"(referring to his answer to Q6).

Stage 3

Measures

One of the aspects that we were interested was to identify if the social mechanisms of memory used in this companion system lead to some level of companionship between user and agent. With MAY we wanted to build a companion that stays with the user for a very long time, developing a relationship and getting to 'know' the user's preferences and wishes [8].

In human relationships it is possible identify a type of relationship that fits in this definition – friendship. Therefore, we employed the McGill Friendship Questionnaire (MFQ) [26], which measures six dimensions of friendship between people. We used this questionnaire to classify the relation between MAY and user (Amy). Mendelson has identified six different functions of friendship conceptually distinct:

1. *Stimulating Companionship* – engagement in enjoyable, fun and exciting activities;
2. *Help* – guidance, aid and other types of help;
3. *Intimacy* – being attentive to other's states and needs and open to honest expressions about thoughts, feelings and personal information;
4. *Reliable Alliance* – keeping available and loyal;
5. *Self-validation* – comfort, encourage and help the friend on keeping positive confidence in himself; and
6. *Emotional Security* – providing comfort and confidence in novel or threatening situations.
7.

According to Allan [1] friendship is "a relational term which signifies something about the quality and character of the relationship involved". Similarly, we wanted to

measure the type of the relationship that the agent is capable of developing with the user based on what it knows about him/her. With MFQ we can somehow get a glimpse of the quality of the relationship between Amy and MAY. In particular, we were interested to examine if positive aspects of the relationship [12] prevail in memory condition in comparison with the no memory condition. The used assertions of MFQ were manipulated to fit into our experiment.

That adaptation was based on the results obtained by [21] in a online survey, in which the users had to associate a set of assertions with the dimensions of MFQ. Our experiment and interactions within both conditions are presented in [6]. Notice that the Emotional Security dimension was not measured in this study. The main reason resides on the fact that in order to recognise this dimension one needs much more than two interactions for the users to understand how MAY can provide comfort to them, as Leite et al [21] pointed in their study. Plus, the agent is only prepared to recognise situations that need a concerned response at the initial state of the dialogue and does not recognize emotional states during the interaction.

Results

At this stage of the evaluation, we applied the MFQ to measure the quality and the characteristics of the relationship. The dimensions *Intimacy* and *Companionship* were the ones that differ significantly ($p < 0.001$) in the extent of both conditions (see fig. 7). The dimensions Help, Self-Validation, Reliable Alliance are embedded in the dialogue structure, so the pro-active feature or 'attitude' of the agent contributes to the high level of agreement with the correspondent statements, in both versions of the questionnaire. The agent encourages the user, gives advice and has an empathetic behaviour that helps in the maintenance of the relationship. As Fig.7 shows, that differences were not significant for those conditions. However, looking at the participants justifications it was evident that they understood the agent's notion of time and that events talked about were theme of conversation between Amy and MAY in some past interaction(s). For example, some users' justifications were "Probably that information was mentioned before"; "Amy told MAY before"; "MAY knew that Amy usually goes to the cinema with April"; "MAY thought that Amy was going with April".

Having or not having memory can explain 11% of the total variability in M7 and 26% of the total variability in M2 (intimacy and companionship dimensions respectively) that suggests a wide effect of our experimental condition in these two dimensions, which we believe to be equally important in the maintenance of a relationship and enough to satisfy hypotheses. Some studies, such as [12] verified that when friends are seen as a primary source of social support, companionship and intimacy are quite important factors that make those relationships endure.

8   Discussion

The overall goal of this last evaluation was to measure the user's perception of the relationship that the system might have developed with the user, based on its memory

functions. The experiment was done in three stages using the persona Amy. Users had to classify the relationship between Amy and MAY, after the agent show some knowledge about her life.

The positive results on *companionship* and *intimacy* dimensions were interesting results for a first prototype. Intimacy is a dimension that is strongly related to long-term, as it develops over several interactions between two people. This may suggest that use of the *persona* Amy allowed us to simulate a long-term interaction in a short duration evaluation, letting us to glimpse, which dimensions would have emerged over time. However, a longitudinal study may reveal other effects on the user than laboratory studies. Certainly, the use of *personas* must not undermine other types of UCD evaluation methods, but rather complement them.It is clear that the described paradigm is adapted to a niche: *companion* systems, which depend on memory and also on several interactions with a user to perform properly. However, is such situation, the use of *personas* seems adequate due to its specific definition and properties, allowing to simulate informed previous interactions with the system.

Friendship Dimensions	Id	Question Statement	Descriptive Statistics		Mann-Whitney differences between conditions
			memory (N=46)	*no memory (N=44)*	
			Mdn[Qaurtiles]	Mdn[Quartiles]	
Companionship	M1	MAY is friendly	5[4,5]	5[4,5]	U = 842.000 $p < 0.05$(1-tailed) r = -0.17
	M2	MAY was companion during interaction	4[4,5]	3[2,4]	U = 433.500 $p < 0,001$ r = -0.512
Help	M3	MAY's comments are useful	3[2,4]	3[2,4]	U = 955.500 *ns*
Self-Validation	M4	MAY does Amy feel that she can do things well	4[3,4]	4[3,4]	U = 975.500 *ns*
	M5	MAY encourage Amy	5[4,5]	5[4,5]	U = 995.500 *ns*
Intimacy	M6	MAY have interest about Amy's life	5[4,5]	5[4,5]	U = 949.000 *ns*
	M7	MAY knows Amy well	4[3,4]	3[3,4]	U = 639.000 $p < 0.05$ r = -0.332
	M8	MAY knows facts about Amy's life	4.5[4,5]	4[2,4]	U = 582.000 $p < 0.001$ r = -0.383
Reliable Alliance	M9	MAY woul still Amy's friend even if a month pass	5[4,5]	5[4,5]	U = 982.500 *ns*

Fig. 5. Mann-Whitney statistics for the agent's characteristics
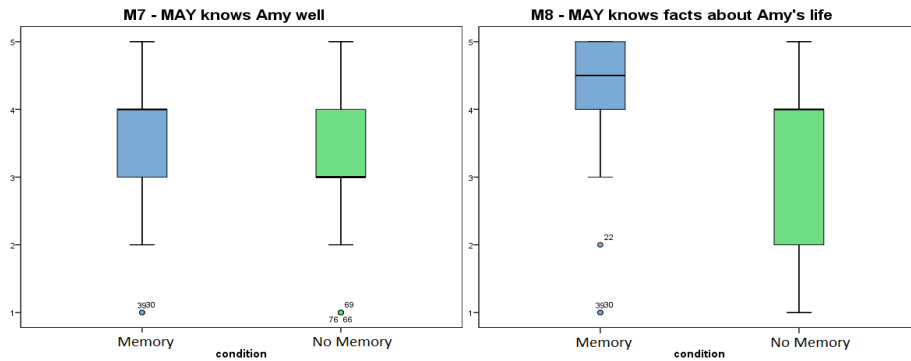
Fig. 8. Differences between the two conditions - *Intimacy* dimension

9   Conclusion

User-centered design (UCD) practices are employed for accessing a clear understanding of the user and task requirements, as well as, giving the user an active role in software development [22]. It is clear that when developing *companion* systems the design must drawn in potential users, as the system should be engaging over long periods of time.

In this paper we described the design cycle of a *companion* system focusing on the essential processes of UCD using the archetype *persona* to guide design decisions. Furthermore, we suggest using the concept of *persona* to the early evaluation stages. As the *persona* is a complete and carefully defined character it would be possible to fill the *companion's* memory with the tailored *personal* experiences. We believe that by following this paradigm we can simulate previous interactions and access the agent's performance, whereas we are simulating a long-term interaction in a short duration evaluation.

References

1. Allan, G., (1989). Friendship: Developing a Sociological Perspective. Westview Press
2. Benyon, D., Mival, O. Landscaping personification technologies: from interactions to relationships. In: *CHI'08 extended abstracts on Human factors in computing systems*.
3. Benyon, D., Bradley, J., Mival, O., Webb N. *Evaluation Paradigm.*
4. Bluck, S. (2003). Autobiographical memory: Exploring its functions in everyday life. Memory 11:2

5. Campos, J., Paiva, A.. "MAY : My Memories Are Yours." In *10th International Conference on Intelligent Virtual Agents*, 406-412. Springer Berlin/Heidelberg, 2010.
6. Campos, J., "MAY : My Memories Are Yours. An interactive companion that saves the user's memories", *Master Thesis,* 2010
7. Castellano G, Aylett, R., Dautenhahn, K., et al. Long-term affect sensitive and socially interactive companions. In: *Proceedings of the 4th International Workshop on Human-Computer Conversation*. 2008
8. Catizone, R., Worgan, S., Wilks, Y., Cheng, W. (2008). A multimodal conversational companion for reminiscing about images. *Fourth International Workshop on Human-Computer Conversation*, Bellagio, Italy
9. Chang, Y., Lim, Y., Stolterman E., (2008). Personas: from theory to practices. In NordiCHI '08: Proceedings of the 5th Nordic conference on Human-computer interaction, New York, NY, USA, pp. 439–442. ACM
10. Chung, C., Pennebaker, J., (2007). Social communication: Frontiers of social psychology, Chapter The psychological functions of function words, pp. 343–359. Psychology Press.
11. Cooper, A. "The Inmates are running the Asylum." In *Software-Ergonomie '99, Design von Informationswelten, Gemeinsame Fachtagung des German Chapter of the ACM, der Gesellschaft für Informatik (GI) und der SAP AG*, 17–. Teubner, 1999.
12. de Souza, L. K. (2006). Amizade em adultos: adaptação e validação dos questionários McGill e um estudo de diferenças de género. Universidade Federal do Rio Grande do Sul.
13. Dix, A., J. Finlay, Abowd,G. D. , Beale, R., (2004). Human-Computer Interaction (3rd ed.). Prentice Hall.
14. Dunbar, R. I. M., Marriott A., Duncan, N. D. C., (1997). Human conversational behavior. Human Nature 8(3), 231–246.
15. Gaver, B., Dunne, T., Pacenti , E., (1999). Design: Cultural probes. Interactions 6(1)
16. Gomes, D. R., Pereira, C. A. , Petreca, T. M. , P. A. Jr. (2008). Teste com usuário baseados em personas em beneficio da prototipaçao em sistemas web. In IADIS Ibero-Americana.
17. Grudin, J. (2006*). In the Persona Lifecycle, keeping people in mind through- out product design*, Chapter Why Personas Work: The psychological evidence, pp. 643–663. Elsevier.
18. Grudin, J., Pruitt, J. (2002). Participatory design and product development: An infrastructure for engagement.
19. Hakulinen, J., M. Turunen, C. Smith, M. Cavazza, Charlton D., (2008). A model for flexible interoperability between dialogue management and domain reasoning for conversational spoken dialogue systems. *Fourth International Workshop on Human-Computer Conversation*, Bellagio, Italy.
20. Leite, I., Mascarenhas, S. , Pereira A., Martinho, C., Prada, R. , Paiva, A.(2010). Why can we be friends? an empathetic game companion for long- term interaction. In Proceedings of the 10th International Conference on Intelligent Virtual Agents.
21. Maguire M. Methods to support human-centred design. *International Journal of Human-Computer Studies*. 2001;55(4):587-634.
22. Picard, R. W. (2000). Affective Computing. MIT Press.
23. Pruitt, J., Grudin, J.,(2003). Personas: practice and theory. In DUX '03: *Proceedings of the 2003 conference on Designing for user experiences*, New York, NY, USA, pp. 1–15
24. Mendelson, M. J. (11 Nov, 2009.). Measuring friendship quality in late adolescents and young adults: Mcgill friendship questionnaires. Canadian Journal of Behavioural Science.
25. Field, A., Hole, G., *How to Design and Report Experiments*, 2003.
26. Mahlke S. Understanding users' experience of interaction. In: *Proceedings of the 2005 annual conference on European association of cognitive ergonomics*. University of Athens;
27. Norman, D. A. (2004). Emotional Design: why we love (or hate) everyday things. NY: Basic Book.