

Designing Friends

Dirk Heylen, Anton Nijholt, Bas Stronks, and Paul van de Vet
Parlevink Research Group
Faculty of Computer Science
University Of Twente, PO Box 217
Enschede, The Netherlands
{heylen,anijholt,vet}@cs.utwente.nl

Abstract. Embodied Conversational Agents are virtual humans that can interact with humans using verbal and non-verbal forms of communication. In most cases, they have been designed for short interactions. This paper asks the question how one would start to design synthetic characters that can become your friends. We look at insights from social psychology and propose a methodology for designing friends.

“The imaginary friends I had as a kid dropped me because their friends thought I didn’t exist.”

Aaron Machado

1 Introduction

This paper is about the design of Embodied Conversational Agents (ECAs). In this field of Human Computer Interaction (HCI) and Artificial Intelligence (AI), the design of ECAs, or ‘virtual humans’, and the communication between those agents and human users is the main object of research. A lot of effort is put into making ECAs more lifelike and believable and making communication with ECAs more effective, efficient, and fun. One way to improve communication in this way is to make the agent more actively involved in building a relationship with the user. From real life we all know that people that you like (or your friends) are able to help you better, teach you better, and generally are more fun to interact with, than people that you don’t like. However, who one likes and why is very much person dependent. In designing agents that can build up a personal relationship this dependence should be taken into account.

These observations sparked our interest in studying the application, effects, and design of a ‘virtual friend’: an agent that is able to observe the user, and with its personality, appearance and behavior is able to respond to the (implicit) likes and dislikes of the user, in such a way that it can become friends with the user and create an affective interpersonal relationship. Such an agent could have additional benefits over a ‘normal’ Embodied Conversational Agent in areas such as computer assisted teaching and entertainment.

We started our research by looking for a sound foundation and not just our lay psychologist’s intuitions) on which to build the design of such a ‘virtual friend’. This led us to investigate the field of psychology. Extensive knowledge about human interpersonal relationships can be found in the

subfields of personality theory and social psychology. Lay psychology and folk wisdom is often proven wrong by psychology research. For example, folk wisdom dictates that, in human interpersonal relationships, ‘opposites attract’. If a researcher would try to increase Human-ECA attraction by designing an ECA that is exactly the opposite of the human in terms of attitudes, it would be likely that the ECA would not be effective, because it has been shown that attraction is greater when these characteristics are similar. Thus, ECA design decisions need to be based on sound psychology research and not on ‘folk wisdom’ or otherwise.

Recent computer science history also shows that insights from psychology can be fruitfully applied in computer science. The application of cognitive psychology in computer science in the last two decades has resulted in improvements in problem-solving skills and task related behavior of computer users, thus improving Human-Computer Interaction effectiveness. Because communication with an ECA is social in nature, improving Human-ECA interaction should be accomplished by applying findings from social psychology. Consequently, a main aspect of our research is the application of (social) psychology in ECA design.

In this paper we first formulate our research questions, showing our research methodology in designing for friendship (Section 2). Next, we consider what we mean by friendship and how we can exploit the tendency we see in humans to assign human properties to animals and objects when we interact with them (Section 3). Section 4 is devoted to the possibility of adopting the main aspects of human-human friendship to human-ECA friendship. In Section 5 we discuss how we can incorporate our findings in the ECA-design process, using a scenario-based design. Tutors would be one of the areas where these theories could be applied (Section

6). We distinguish between the initial design of an ECA and the possibility to change the ECA characteristics according to an adaptation strategy based on knowledge obtained by interacting with a particular user in Section 7.

2 Research Questions

The aim of our research is to extend the knowledge in the field of Embodied Conversational Agent technology in Human-Computer Interaction, by looking in more detail at human-ECA relationships and, specifically, human-ECA friendship. The general goal is to make Human-Computer Interaction more effective. The research questions guiding this research are threefold.

The first question we have asked ourselves is “What is known about friendship in psychology research?” In answering this question, we explored the psychology behind friendship, including:

- What are the psychological mechanisms behind the formation of friendship?
- What are the variables and issues affecting friendship?
- What are the effects of friendship on interaction?

The second question concerns the way we could apply the answers to the first question to Human-Computer Interaction. Answering this question, means that we have to find a way to take the findings from psychology and transfer them to the human-ECA situation. Some of the subquestions here are:

- Is it possible to apply social psychology to Human - ECA relationships?
- How can we translate the knowledge about friendship in social psychology to useful insights in human-ECA friendship?

Finally, we have to think about how we could go about designing and creating virtual friends. The third research question is concerned with applying the translation of human psychology to the design phase, design issues, and architecture of ECAs. Some of the more specific questions to be answered here are:

- Given the translation of the psychological insights, how can the design phase of an ECA be adapted to accommodate the creation of a ‘friendlier’ ECA?
- What are important issues in the design of friendly ECAs and how can they be addressed?
- Given what is known about human-ECA relations, how can the architecture of an ECA be adapted in order to create friendlier ECAs?

Because of the novelty of the object of research, this research has an exploratory and theoretical

character, and does not focus yet on implementation. In this paper, we will also not specify answers to all these questions in detail. For more information, we refer to [21]. The main way to find answers to the first research question has been by reviewing the psychological literature on friendship. Specifically research in social psychology on interpersonal relationships and attraction has been of interest. A summary of the relevant theories is presented in the next section.

3 On Friendship

Interpersonal relationships are the subject of research in social psychology. There are several theories that are concerned with explaining why people start and maintain relationships. We briefly review some of them to introduce some of the issues involved in friendships.

In reinforcement theory, for instance, it is assumed that we are attracted to a person when we are rewarded in the presence of that person. Rewards don’t have to be material, of course. If the other confirms our attitudes for example, this can also qualify as a reward. Byrne’s ‘Law of Attraction’ makes this idea more specific [4]. It says that the attraction towards a person A is a positive linear function of the proportion of positive reinforcements (positive reinforcement divided by total reinforcements) received from A .

Rewards are also central to social exchange theory. It is assumed here that for attraction and friendship rewards must outweigh the costs. Commitment to a relation relies on investments, rewards and available alternatives of the relation.

In equity theory it is assumed that fairness is a central issue in relationships. The perceived input/outcome ratios of both partners involved should be about equal.

Friendship is a specific kind of interpersonal relationship. An accepted definition (cited in [11]) is the following.

[Friendship is] a voluntary interdependence between two persons over time, that is intended to facilitate social emotional goals of the participants, and may involve varying types and degrees of companionship, intimacy, affection, and mutual assistance.

Various stages can be distinguished. In the initiation phase an assessment of attraction takes place. Some attributes may lead to the immediate exclusion of people as friends, whereas others make it less likely that the other is considered as a potential friend. The latter attributes function as disregard criteria: for instance, one would not immediately consider much older or much younger people as candidates for friendship. Once friendship is established, a maintenance phase is set in, involving, amongst

others, conversations about attitudes and personal issues. A willingness for self-disclosure (including how one feels) is an important prerequisite for friendship. Friendships may terminate due to negative changes in the relationship.

There are many factors that play a role in the way friendships are set up and evolve such as gender, age, social class and ethnic background. For instance, it has been said that male friendships are activity-based (doing things together), whereas female friendships are based on affect.

Interpersonal attraction is an important factor in friendship. It is governed by positive reinforcements (cf. 'Law of Attraction') and similarity between subjects is also a key factor. To put it simple, we don't dislike people that are just like us. Hence, similarity of attitudes, personality, ethnicity, social class, humor, etc., reinforces the relationship. Other issues that play a role are physical attractiveness (the 'halo effect') and reciprocity of liking (whether we think that the other person likes us). The perceived reciprocity of liking is especially important in the initial stage of friendship formation.

Friendship has a number of effects on interaction. This includes an increase of altruistic behavior. It can have a positive impact on task performance (and attribution of task success).

Summarising, some of the elements that play a crucial role in the establishment and maintenance of friendship relations are: similarity of attitudes, the notion of rewards, joint activities or relational values (trust, confidence, intimacy), reciprocity, certain aspects like social class and common interests that define exclusion or disregard criteria.

4 Translating Friendship

How do we translate the issues that play a role in human-human friendship to issues that can be implemented in human-ECA friendship? One place to start is with the studies about 'Computers Are Social Actors' (CASA) [19] in which experimental evidence is gathered to support the so-called *Media Equation*: media equals real life or, people respond to mediated worlds as if they were real. More particularly, in interactions with a computer, people treat the computer similar to another human being. This can be viewed as a specialization of the anthropomorphic tendency of humans. Given minimal cues, people attribute personality and gender stereotypes to computers, or respond to automated flattery as if it were given by a human. The research also shows that users do not believe the computers *are* human, nor that their complexity makes them *like* humans. The results indicate that users *respond* to computers *as if* they were humans, an important finding for our research. Of course, this doesn't mean that people will interact with computers *exactly* as they do with humans.

Shechtman [20] conducted experiments to study relationship behavior during keyboard human-computer interaction and (apparently) keyboard mediated human-human interaction. In the latter case participants used much more communion and agency relationship statements, used more words and spent more time in conversation.

It must also be noted that in the CASA experiments, the cues that were used to elicit the anthropomorphic responses were minimal. Word choice, for instance, elicited personality attribution, voice pitch elicited gender attribution. In ECAs, however, the cues are not even minimal. Gender can be communicated by means of physical appearance and voice, personality can be communicated by means of behavior, word choice and nonverbal communication, much like is the case in human-human interaction. Consequently, the CASA paradigm should be applicable to ECAs at least as well as to computers in general.

Hence, users will attribute gender, attitudes, personality, emotion, ethnicity and other human characteristics to ECAs. These issues are instrumental in friendship formation between humans and therefore translation to the human-ECA situation will make it possible to generate human-ECA friendship.

One of the central issues concerning friendship is the notion of reward as discussed by a number of theories. A way to create friendship in the human-ECA setting is to look for a possible implementation of reinforcement theory. From that perspective, the way to create a relationship between a human and an agent is to provide the human with rewards. These rewards can be numerous: fun, information, or other positive reinforcements. Hypothetically, if we use positive reinforcement, the human will become attracted to the agent.

Applying Social exchange theory suggests that the rewards of a relationship must outweigh the costs. Hypothetically, the same general principle can be applied to human-ECA relationships: rewards for the human need to be higher than costs for the human. Examples of rewards for humans are fun interactions or positive reinforcements, examples of costs are irritation or negative reinforcements.

Equity theory suggests that the perceived input/outcome ratios of the parties involved should be equal. Hypothetically, this should be true for the human-ECA situation as well. This means that the ECA should not be the only one to contribute towards a relationship, but the ECA should also 'expect' contributions from the user and 'gain' something out of the relationship. Communicating reciprocity of liking and letting the human show that the ECA appreciates the relationship could be an application of this theory.

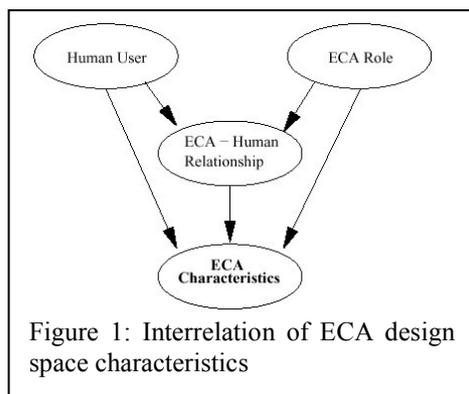
In order to have a friendship relation with a human, the ECA also needs to obey the rules of human–friendship relations that deal with time. The friendship relationship needs to be initiated and maintained, and in each stage the ECA needs to conform to the rules that govern these stages of friendship, including a stage of decline. This requires different behavior during the different stages, as explained in the previous section.

Further considerations about the translation of mechanisms involved in friendship formation to the human–ECA situation include examining the influence of certain human characteristics. For example, when users are younger (or males) they may base the relationship with the ECA more on commonality, whereas older users (or females) may deem the relational features of a friendship as more important. As discussed in Section 4, sexual attraction, social class and age similarity, and ethnicity also influence attraction and friendship and need to be considered.

5 The Design Phase

An ECA can be described on different levels. On the *action and interaction* level, the ECA is described as a collaborator with a domain role, an interpersonal role and competence in a number of areas.

In the design decisions about the characteristics and abilities of the ECA, one should deal with the factors affecting friendship by considering also the user characteristics, the role the ECA plays in the domain, and the intended personal role (Figure 1). These involve decisions on both the action and interaction level. For the characteristics we need to be aware of the stereotypes attached to them.



Areas of competence include domain competence (the necessary, sufficient and desirable domain knowledge and competence an ECA should possess in order to be perceived competent in its domain role), the relational competence (social knowledge and competence an ECA should possess in order to

be perceived competent in its interpersonal role) and communication competence (issues regarding the conveying of domain role and relational information). Obviously, the relational competence is dependent on the intensity of the desired human–ECA relation. An example of a relational ability is the degree of user modeling that the ECA is able to do in order to ‘get to know’ the user. An example of a communication skill (both verbal and nonverbal) related to the interpersonal role of the ECA is the ability to communicate reciprocity of liking. In general communication skills are very important. One should just consider the case of nonverbal communication, in this matter. Facial expressions even of an emotional character are not just displays of what a person feels, but in most cases, about a feeling a person wants to communicate consciously, as has been pointed out repeatedly (consider for instance the case of smiling and laughter as described by Kraut and Johnston [16]; they present a clear case where people smile not because they are happy, but because they want to “show” they are happy). As Argyle discusses in his book on interpersonal behavior [1], many facets of the relationship between humans influence the way they communicate nonverbally: physical proximity, orientation in conversations, bodily posture, gestures, head nods, facial expressions, eye-movements, all relate to values on interpersonal scales such as intimacy and dominance, attitudes towards the other (cooperative, competitive,...), tenseness and relaxation, anxiety and trust, reinforcements (rewards and encouragements), attention, intimacy, etcetera.

In order to decide about the characteristics and abilities a high abstraction level ECA design tool is needed, a *design method for the character of an ECA*. This design method has to facilitate consideration of the interrelation of user characteristics, ECA characteristics and roles in order to increase the likelihood of friendship formation. We propose a design process that is based on general scenario-based design and a particular specialization presented in [9] with a focus on the earlier mentioned friendship aspects of ECA design.

Scenario-based character development, adapted from character design in the dramatic arts, is an iterative prototyping method using scenarios, story boarding, and group techniques. In the process, the developer collaborates with the user or users to build a model of the ECA’s character by means of scenarios, storyboards, and sample interactions, that enable the user and the developer to envision the ECA in its future role. The aim of the character to be designed is to be tailored to the specific user group characteristics in order to increase the friendship likelihood. This is the reason why future users and social psychology experts play an important role in the design process.

There are several reasons why scenario-based design is practical as the design process of ECA characters. First, scenario-based design has already been used successfully in developing ECAs, so its validity is already shown. Furthermore, this application of scenario-based design in ECA design showed that the method is useful for putting pieces of different research in ECA technology together. Most importantly, the design method is based on designing explicitly with social interaction in mind, and can be used to focus on the ‘soft’ side of the ECA design, the personality, roles, and abilities.

6 Potential Applications

So far, not many embodied conversational agents have been designed and implemented with friendship in mind. There are several factors that make it hard to build synthetic friends. Many existing ECAs are used for superficial, one-time interactions. Friendship, however, is something that needs time to develop and evolve. Most ECAs are also made with a specific task in mind: to provide information or to entertain. These tasks are not always of the kind that friends would typically engage in. However, one of the other common tasks that ECAs have been used for, namely tutoring, might qualify as a good candidate for a domain to invest in when designing agents that can build up friendship relations with their students. In this case a lot of things that friends are expected to provide for one another ([10]) are present. Depending on the kind of tutoring, students might be prompted to show quite a good deal about themselves and their attitudes. They will at least reveal how well they do at the specific task. In a personalised approach, they will want to present a positive image of themselves. They will want to perform well and be rewarded for that. But also, a tutor should invest in building up rapport, motivate them to do things, support, and help them, taking emotional care if the learning is not that easy, praising them if things go well, or making the lessons fun. This task involves a certain degree of intimacy and requires a high degree of trust. The tutoring situation also leaves room for the student to reward the tutor: just by trying harder, by succeeding or by explicit praise. The tutor could also be cast as a student’s peer and lessons could take on the form of solving problems or doing homework together. Having someone from a similar age to help you can make it more likely that the ECA is considered suitable as a candidate friend (cf. disregard criteria). Tutoring ECAs can also be involved in teaching sessions that are spread over longer time periods. In this light, it is interesting to look at Bickmore’s recent study of the exercise advisor agent Laura as this is the first attempt to implement and evaluate a system in which users interact with a personal coach for a longer period of time [3].

7 Adapting to the Human Partner

“.. and comes face to face with a receptionist [avatar]. For a moment, he can't peg her racial background; then he realizes that this [avatar] is half-black, half-Asian – just like him. If a white man had stepped off the elevator, she probably would have been a blonde. A Nipponese businessman would have come face to face with a perky Nipponese office girl.”

Snow Crash, Neal Stephenson, 1992

Another way to increase the likelihood of friendship is to align the characteristics of the user and the ECA during interaction. Friendship and attraction can be facilitated by *adapting the character* of the ECA to the perceived characteristics of the user (see Figure 2). The *ECA characteristics updating* part of the ECA changes the ECA characteristics according to some adaptation strategy. This strategy should be directed at finding the optimal set of characteristics for friendship, given a user model. Some form of user modeling should enable the ECA to extract the user personality and other characteristics from the user’s input. Because the general tendency of humans is to like people that are similar to them, a plausible strategy could focus on adopting characteristics (personality, attitudes) that are similar to those of the user.

Not all characteristics can be adapted. Certain issues in the environment of the ECA limit the adaptable characteristics. For example, if the personality of a sales-ECA would be adapted to be completely similar to the personality of a user, problems will arise when an extremely introvert user uses the agent. Constraints like this need to be considered in the design phase of the agent, and integrated in the adaptation strategy of the ECA.

Adapting the computer’s personality has been shown in user studies in a laboratory setting [19] to increase perceived liking for a computer that is communicating a certain ‘personality’ with minimal (pre-programmed) cues. The studies showed that when a user was an introvert, and the computer changed its personality from extravert to introvert during the interaction, the perceived liking increased. This was also true for an extravert user and a computer personality changing from introvert to extravert.

The idea of adapting personality by some strategy is also discussed in the ‘Peedy’ project [2]. With the Peedy ECA, the emotional state and personality of the user are assessed using Bayesian belief networks. After assessing the state of the user, the emotional state and personality of the ECA can be changed according to some strategy. Then, theoretically, the psychological state of the ECA can be updated, partly by using the same Bayesian belief networks.

Although in the Relational Agent project at the MIT [5] the ECA does not adapt to the user, the results from the user study indicate that adapting ECA behavior to the user personality could be effective. A study showed that extravert users liked a small-talking ECA better than a non-small-taking ECA, whereas introvert users liked a non-small-talking ECA better than a small-talking ECA. Obviously, considering small talk as a behavior that is exhibited more by extravert humans than by introvert humans, this liking is explained by our previous observations. If the personality of the user would be perceived by the ECA in one way or the other, the communicated personality of the ECA (by use of small talk) could be adapted to be similar to the personality of the user. This way, the extravert user as well as the introvert user would be optimally attracted to the ECA.

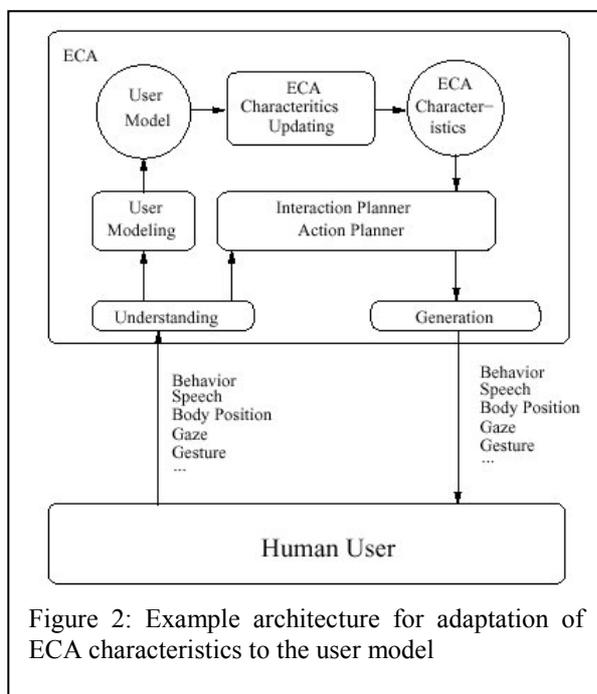


Figure 2: Example architecture for adaptation of ECA characteristics to the user model

Conclusions

In this paper we have confined ourselves to a theoretical research perspective. Neither the design method proposed nor the adaptive architecture could be applied or tested. In order to test the scenario-based character design method ECA characters should be developed using his method. Depending on the experiences of the design process, the proposed method should be further refined and adapted. Research in adaptive ECA architectures should be continued by implementing an ECA architecture that adapts to the user, based on the technique proposed in our research.

In our previous research, we have introduced several ECAs with varying degrees of verbal and nonverbal communication skills ([6], [13,

14,15,17,18]). Our research programme also includes designing agents that have personality and show emotions [7]. In this paper we made a first step towards introducing ECAs that work on the formation of long-term relationships with their human conversational partner. Implementation and evaluation are the necessary next steps.

References

- [1] Argyle, M. *The Psychology of Interpersonal Behavior*. Penguin Books, London (1972).
- [2] Ball, G. and Breese J.: "Emotion and personality in a conversational agent"; In: [8].
- [3] Bickmore, T.W.: *Relational Agents: Effecting Change through Human-Computer Relationships* PhD, MIT (2003).
- [4] Byrne D.: "Attitudes and attraction"; *Adv. in Experimental Social Psychology*, 4: pp.34-89 (1969).
- [5] Bickmore, T. and Cassell J.: "Relational agents: A model and implementation of building user trust"; *ACM CHI 2001 Conference Proceedings*, Seattle, Washington; (2001).
- [6] The Bui, D. et al.: "ParleE: An adaptive plan based event appraisal model of emotions". Parlevink Internal Report, University of Twente, April (2002).
- [7] Carroll, J.M. (ed.): *Scenario Based Design: Envisioning Work and Technology in System Development*. John Wiley, New York (1995).
- [8] Cassell, J. et al. (eds.): *Embodied Conversational Agents*. MIT Press, Cambridge (2000).
- [9] Churchill E.F.: "May I help you? Designing embodied conversational agent allies" In: [8].
- [10] Duck, S. *Meaningful Relationships*. New York: Guilford Press (1991).
- [11] Fehr, B: *Friendship Processes*. SAGE Publications, Thousand Oaks, California, (1996).
- [12] Friedman, B (ed.): *Human Values and the Design of Computer Technology*. Cambridge University Press, Stanford, (1997).
- [13] Heylen, D. and Nijholt, A.: "Designing and implementing embodied agents: Learning from experience." *Proc. Multimodal Communication and Context in Embodied Agents*. C. Pelachaud & I. Poggi (eds.), Workshop at 5th Intern. Conf. on *Autonomous Agents (Agents 2001)*. Montreal, 77-82 (2001). (*CEUR Workshop Proc.* Vol-46, 2001; <http://ceur-ws.org>.)

- [14] Heylen, D., Nijholt, A. and Poel, M: "Embodied agents in virtual environments: The Aveiro project"; *European Symp. on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems*, Tenerife, Verlag Mainz, Aachen, 110-111 (2001).
- [15] Heylen, D. et al.: "Gaze behavior of talking faces makes a difference."; *Proc. ACM-CHI 2002*, Minneapolis, April (2002).
- [16] Kraut, R. E., Johnston, R. E.: "Social and emotional messages of smiling: An ethological approach." *Journal of Personality and Social Psychology*, 37, pp.1539-1553; (1979).
- [17] Nijholt, A.: "From virtual environment to virtual community" *New Frontiers in Artificial Intelligence*. T. Terano et al. (eds.). Lecture Notes in Computer Science 2253, Springer Verlag, Tokyo, pp.19-26; (2001).
- [18] Nijholt, A.: "Agents, believability and embodiment in advanced learning environments."; *Proc. IEEE Intern. Conf. on Advanced Learning Technologies (ICALT 2001)*, T. Okamoto et al. (eds.), Madison, IEEE, Los Alamitos, 457-459; (2001).
- [19] Reeves, B. and Nass, C.: *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press, (1996).
- [20] Shechtman, N. and Horowitz, L.M.: "Media inequality in conversation: how people behave differently when interacting with computers and people" *SIGCHI-ACM CHI 2003: New Horizons*, ACM, New York, pp. 281-288 (2003).
- [21] Stronks, B.: "Friendship relations with embodied conversational agents: Integrating social psychology in ECA design". Report, Parlevink Research Group, University of Twente, (2002).