The Persona Zero-Effect: Evaluating virtual character benefits on a learning task with repeated interactions

Jan Miksatko[1], Kerstin H. Kipp[2], and Michael Kipp[1]

[1] DFKI, Embodied Agents Research Group, Germany
{jan.miksatko,michael.kipp}@dfki.de
[2] Saarland University, Experimental Neuropsychology Unit, Germany,
k.kipp@mx.uni-saarland.de

Abstract. Embodied agents have the potential to become a highly natural human-computer interaction device – they are already is use as tutors, presenters and assistants. However, it remains an open question whether adding an agent to an application has a measurable impact, positive or negative, in terms of motivation and learning performance. Prior studies are very diverse with respect to design, statistical power and outcome; and repeated interactions are rarely considered. We present a controlled user study of a vocabulary trainer application that evaluates the effect on motivation and learning performance. Subjects interacted either with a no-agent and with-agent version in a between-subjects design in repeated sessions. As opposed to prior work (e.g. Persona Effect), we found neither positive nor negative effects on motivation and learning performance, i.e. a *Persona Zero-Effect.* This means that adding an agent does not benefit the performance but also, does not distract.

Key words: Embodied conversational agents, human-computer interaction, computer supported learning

1 Introduction

Embodied agents can be a powerful user interface because they are capable of *multimodal interaction*: they can communicate with verbal and non-verbal channels, react to user behavior and give subtle feedback in the form of a smile or a nod. Moreover, embodied agents can engage users in a *social* way, becoming supervisor, audience or virtual friend, and thus increasing commitment, motivation and ultimately performance. Applications of embodied agents include virtual teachers, presenters in museums or personal shopping assistants. In the domain of education, virtual characters have received special attention because their potentials in terms of nonverbal communication and social relations, both important aspects of a good *human* teacher.

The benefits of embodied agent presence in computer based learning environments have been analyzed in multiple prior studies (cf. [1, 2]). Generally, an embodied agent allows for a *richer interaction* since the communication is multimodal (gaze, facial expressions, head nods and gestures) [3]. An important question is whether the agent can *motivate the student to learn longer* with the system [4, 5]. Also, a positive effect on *learning performance* (e.g. recall, comprehension and problem solving) has been suggested [6, 7]. A learning effect can be due to the agents' *stimulation of the learning process* by guiding the user, directing attention, and encouraging exploration and reflection [8, 9] and also may benefit from a *personal relationship* similar to that between teacher and student [10]. A related finding is that the agent's presence makes the task and learning material to be *perceived as being easier* than without an agent [11].

The actual findings, however, do not conclusively answer the above questions and are often contradictory. Several studies [6, 7] show learning improvement when an agent is present, whereas other studies [11, 5] show the contrary. The studies not always include a control *no-agent* condition, for instance the Persona Effect study [4, 6] compares only five different agent versions among each other. In terms of study duration, almost all studies assessed their systems in a single session, with the notable exception of Bickmore [10]. However, it could have a decisive impact on e.g. learning performance if the user feels annoyed by the agent after a few sessions or, on the contrary, becomes more comfortable over time. Furthermore, the empirical studies vary significantly in the interaction style (e.g., an agent presenting a text [11, 7] vs. agent giving hints and commenting on user actions [6, 5]), in the agent appearance (a cartoon-style full-body agent [11, 6] vs. a talking head [7, 5]), in the learning task (a problem solving [6, 5] or a memory task [11, 7]) and also in the measures (questionnaires for rating subjective perception vs. quantitative measures for collecting performance data).

We present a user study with a vocabulary trainer application with repeated interactions that evaluates the effect of an agent on motivation and memory performance. The system consists of a *no-agent* version (control condition) and a *with-agent* version. In a between-subjects design, 36 subjects interacted with one system version each. We evaluate the question of motivation and performance in a clean experimental design. In particular, we examine the development over multiple sessions and employ quantitative measures for user motivation. The results show neither positive nor negative effect on the motivation and memory performance. In other words, the agent is not detrimental to learning, but also, does not improve it.

In summary, the major aspects of this work are as follows:

- *Clean experimental manipulation:* We designed a system with two conditions (no-agent, with-agent). The with-agent condition features very few nonverbal behaviors so that the main difference between conditions is the presence/absence of a (moving) agent, following Dehn et al [2] to not introduce confounding aspects between conditions. We used objective measures (e.g. subjective perception of time) for assessing effects of agent presence.
- *Repeated interactions:* Following criticism on existing studies [1, 2], we designed our study to encompass repeated interactions over a 8-day period to emulate how our system is perceived as if in real-world usage.

– *Persona Zero-Effect:* We show that the presence of our agent does not change learning performance or motivation compared with the no-agent system.

2 User study

To investigate the potential benefits of embodied agents in the learning environments we created the ITeach environment [12]. Regarding the question of how the presence of a virtual character influences memory performance and motivation, we have the following hypotheses:

H1 The agent presence has no positive effect on memory performance [11, 5].
H2 The presence of our agent increases the motivation of the learner [9, 4].

2.1 Scenario: A Virtual Vocabulary Trainer

In the present study we used ITeach for a simple vocabulary trainer, based on the flash card learning system. The flash card system is a question/answer game: a card with a foreign language expression is presented and the user answers for herself what the correct translation is. Then, the correct answer is displayed and the user rates: "I knew it" or "I did not know it". All cards are stored in *bin 1* at first. Depending on the rating, a card is moved one bin forward ("I knew it") or backward ("I did not know it"). This means that (1) difficult words are repeated more often than easy words, and (2) the user can easily see progress by looking at the bins where the rightmost one represents "long-term memory".

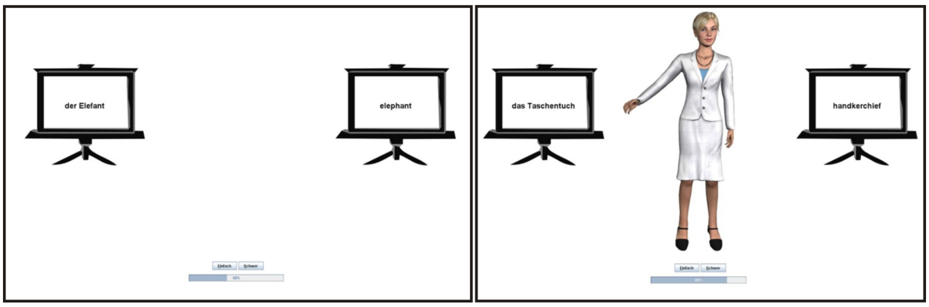

Fig. 1. Screenshots of no-agent and with-agent version of the application.

Each learning session has three phases. First, a fixed number of new unseen cards are presented. Only if such a card is unknown, it is is moved to Bin 1. The goal of this filtration phase is to introduce the new learning material and align the user's knowledge to the same level. Second, the cards in Bin 3, Bin 2 and Bin 1 are presented in the mentioned order. These bins contain material to be reviewed (cards from the previous learning sessions and newly added cards). In the last phase, the user can learn optional vocabulary cards. The session serves as motivation measures discussed in the following section. We selected vocabulary topics of common interest for all users: animals, clothes, body and medicine. The algorithm orders the cards by the topic when selecting them from the current bin.

The user interface was deliberately kept simple (see Fig. 1). In the *no-agent* version, it consists of two windows displaying the English and German expressions and a row of buttons for showing and rating the answer. Each expression is displayed and spoken by a commercial text-to-speech (TTS) engine (Nuance RealSpeak Solo). The system gives positive feedback, e.g. "good!", for known cards 20% of the time. In the *with-agent* version, a female embodied agent is added in the middle of the screen. It features some idle movement to make the agent look alive and does a minimum amount of gestures: pointing to the card 15% of the time and accompanying positive feedback with a smile or nod. Subjects were seated in a cubicle and wore headphones to exclude any external distractions.

2.2 Study Design

In our study we compared two conditions: no-agent and with-agent in a between-subjects design, i.e. the subjects were split into two groups and each group interacted with one condition. The second major design decision was to repeatedly (four times) interact with the system over a longer period of time (8 days). Subjects were briefed before the first session and, during every session, saw short on-screen instructions before the each phase (filtration, learning, open-ended). 36 subjects participated in the study, all students of Saarland University and native German speakers. They were paid 20 EUR for the experiment. In both experiment groups, there were 50% male and 50% female, the average age was 26 years, one third of participants were students of computer science related disciplines and less than 10% of participants indicated greater experience with e-learning systems and virtual agents. We used the following measures for evaluating the agent's influence on learning.

Memory performance measures To measure learning performance, we introduce the measure of *card score* and *bin score*. The card score captures the learning progress within a session by summing up card "movement" between bins. We sum up +1 point for a known card and -1 point for an unknown card (i.e. a movement forwards and a movement backwards respectively). The bin score reflects the overall knowledge as a weighted sum of cards in each bin with weights 0 to 3 for bins 1 to 4, respectively. The higher the bin number the more deeply the contained vocabulary is stored in memory, therefore the higher weight. Both measures are normalized by the total number of total cards.

Motivation measures We employ two motivation measures. The learning phase is split into *legs*: segments of 60-70 seconds length each. After each leg, a *post-leg questionnaire* is displayed. It asks the subject "How long do you think the last leg took?" and provides a slider for giving the time estimate. We hypothesize that if the user is well motivated and engaged, she will estimate the time below the actual time, whereas if the user is less motivated, the subjective time passes more slowly and a higher estimate will be given. We refer to this measure as *perception of time*. The second measure, *optional vocabulary cards*, is obtained in the open-ended phase: the subject is asked, in regular intervals of 5 cards, whether 5 additional cards should be learned. This goes on until the subject decides to quit. The number of additional cards is recorded as a measure of motivation.

3 Results

For all data, 36 subjects, 4 sessions each, we performed ANOVAs with the two factors *agent* (no-agent, with-agent) and *session* (session 1–4). We only report main effects on interactions with factors agent since only the influence of the agent is of interest.

For memory performance we compared card score (learning quality within a session) and bin score (overall knowledge). Fig. 2 shows that our measures indeed reflect the learning progress across the four sessions. As for the two conditions, no-agent and with-agent, the score development is statistically equal when computing a two-factored ANOVA (F(1,3)=.35; p=.79).

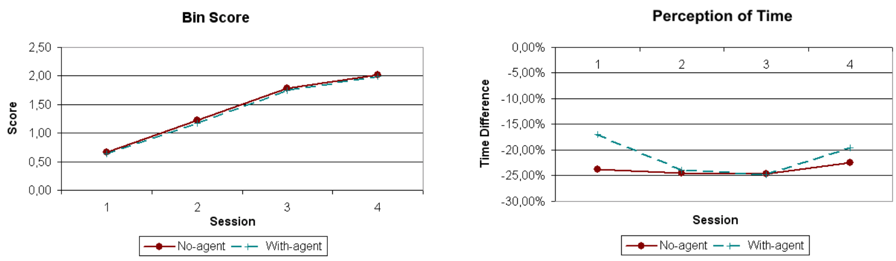

Fig. 2. Bin Score measure for learning performance and Perception of Time measure for motivation.

Motivation was measured using a post-leg perception of time and the number of additional cards. The first measure was analyzed across sessions (see Fig. 2) and across legs (we took the average of the first four legs). An ANOVA with factors agent and session product no main effect (F(1,3)=.58; p=.63). An ANOVA with factors agent and leg yielded a main effect (F(1,3)=.03; p<.05) which was, however, not confirmed in post-hoc analysis. Therefore, we can regard the legwise development as equal. The second motivation measure, the number of additional cards, was also compared between the no-agent and with-agent conditions but yielded no significant result in a two-factored ANOVA (F(1,3)=1.09; p=.36).

4 Discussion

With respect to the memory performance, the users clearly improved their vocabulary knowledge according to the both measures which confirms the validity of our learning setup. However, the measured values are statistically equal in both conditions (no-agent, with-agent). This *confirms hypothesis H1* that the presence of the agent is not detrimental to learning performance — neither is it beneficial. This confirms findings by [11, 5] and contradicts studies claiming to find a positive effect [6, 7]. Note that all mentioned studies only looked at a single interaction, so this is the first study were this zero-effect was found in repeated interactions. There was also no difference in terms of motivation between the no-agent and with-agent conditions. There was neither increase nor decrease in the motivation rating, the users perceived time equally in both versions and did not interact longer or shorter with the system in the open-ended phase. This result *refutes hypothesis H2* that the presence of an agent motivates the learner. It contradicts the finding by Höök et al. [9] where users interacted longer in the with-agent condition and partially supports the results by Bickmore [10, 13] that the number of user logins was not different among the conditions. Note that the interaction sessions were short (about 15 minutes) and the participants were not overloaded with new vocabulary.

5 Conclusions

We presented a user study evaluating the presence of an agent concerning motivation, memory performance and agent perception. In a vocabulary learning task, 36 participants participated in a study where half of the subjects interacted with a system with an agent and the other half with a plain system (no agent). Both systems had the same voice output, so only differed in terms of presence/absence of the agent's body. Memory performance was assessed with quantitative measures. For measuring motivation we assessed the subjects' perception of time and the number of voluntary additional cards. Each subject interacted in four sessions over a period of eight days. Our results showed that the agent presence has neither negative nor positive effect on the memory performance and motivation as suggested by previous single-interaction studies (e.g. [11]) and as opposed to some studies (e.g. [7]). We conclude that adding an agent is not detrimental but also does not beneficial in terms of learning or motivation. This finding can only be generalized with caution, however, since (a) introducing new features may induce distraction and/or a positive learning effect, (b) our finding may be specific to the domain and/or to the particular agent employed.

Our study establishes a solid baseline for future repeated interactions studies. Future studies should identify which kind of additional value is required to induce a positive learning effect in terms of performance and/or motivation (e.g. trying to build a personal relationship with the user, richer multimodal interaction, more learning-related contributions by the agent). An equally important question is which additional features would cause distraction and how to balance distracting and motivating factors.

Acknowledgments. This research has been carried out within the framework of the Excellence Cluster Multimodal Computing and Interaction (MMCI), sponsored by the German Research Foundation (DFG). We would also like to thank to Charamel GmbH for providing us with their Avatar engine.

References

1. Gulz, A.: Benefits of virtual characters in computer based learning environments: Claims and evidence. International Journal of Artificial Intelligence in Education **14**(3) (2004) 313–334
2. Dehn, D.M., van Mulken, S.: The impact of animated interface agents: a review of empirical research. International Journal of Human-Computers Studies **52**(1) (2000) 1–22
3. Cassell, J., Thorisson, K.R.: The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. Applied Artificial Intelligence **13** (June 1999) 519–538
4. Lester, J.C., Converse, S.A., Kahler, S.E., Barlow, S.T., Stone, B.A., Bhogal, R.S.: The persona effect: affective impact of animated pedagogical agents. In: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM New York, NY, USA (1997) 359–366
5. Moundridou, M., Virvou, M.: Evaluating the persona effect of an interface agent in a tutoring system. Journal of Computer Assisted Learning **18** (September 2002) 253–261
6. Lester, J.C., Converse, S.A., Stone, B.A., Kahler, S.E., Barlow, S.T.: Animated pedagogical agents and problem-solving effectiveness: A large-scale empirical evaluation. In: Proceedings of 8th Conference on Artificial Intelligence in Education, Press (1997)
7. Beun, R.J., Vos, E.D., Witteman, C.: Embodied conversational agents: effects on memory performance and anthropomorphisation. In: Intelligent virtual agents: 4th International Workshop, IVA 2003, Kloster Irsee, Germany, September 15-17, 2003: proceedings. (2003) 315
8. Rickel, J., Johnson, W.L.: Animated agents for procedural training in virtual reality: Perception, cognition, and motor control. Applied Artificial Intelligence **13** (1999) 343—382
9. Höök, K., Persson, P., Sjölinder, M.: Evaluating users' experience of a character-enhanced information space. AI Commun. **13**(3) (2000) 195–212
10. Bickmore, T.W.: Relational agents: Effecting change through human-computer relationships. PhD thesis, Massachusetts Institute of Technology (2003)
11. Mulken, S.V., André, E., Müller, J.: The persona effect: how substantial is it? People and Computers (1998) 53–66
12. Miksatko, J., Kipp, M.: Hybrid control for embodied agents applications. In Mertsching, B., ed.: Proc. of the 32nd Annual Conference on Artificial Intelligence. (2009)
13. Bickmore, T.W., Picard, R.W.: Establishing and maintaining long-term human-computer relationships. ACM Trans. Comput.-Hum. Interact. **12**(2) (2005) 293–327