Know Your Users! Empirical Results for Tailoring an Agent´s Nonverbal Behavior to Different User Groups

Nicole C. Krämer[1], Laura Hoffmann[1], and Stefan Kopp[2]

[1] University of Duisburg-Essen, Forsthausweg 2, 47048 Duisburg, Germany
[2] CITEC - Center of Excellence "Cognitive Interaction Technology", Bielefeld University, P.O. Box 10 01 31, D-33501 Bielefeld
{Nicole.Kraemer,Laura.Hoffmann}@uni-due.de,
skopp@techfak.uni-bielefeld.de

Abstract. Since embodied agents are considered as equally usable by all kinds of users, not much attention has been paid to the influence of users´ attributes on the evaluation of agents in general and their (nonverbal) behaviour in particular. Here, we present evidence from three empirical studies with the agent Max, which focus on the effects of participants` gender, age and computer literacy. The results show that all three attributes have an influence on the feelings of the participants during their interaction with Max, on the evaluation of Max, as well as on the participants` nonverbal behavior.

Keywords: virtual agents, nonverbal behavior, user groups, gender, elderly users, computer literacy.

1 Introduction

Within human-computer interaction it has long been proclaimed that due to the fact that humans show inter-individual differences, it is important to "know the user" [1] when designing an interface. In research on embodied conversational agents this goal has not received much attention. This is perfectly plausible, given that an interface agent has always been assumed to be appealing to all kinds of users [2]. In fact, the vision of embodied interface agents as an interface of the future has always been connected to the notion that due to its ability to produce and understand verbal and nonverbal human communication, a tailoring to specific user will not be necessary [3]. At best, the question of appearance, that is, how the agent should look like when being employed in different contexts and for different user groups has received some attention [4]. However, it is not only appearance but also the (nonverbal) behaviour of the agent that might be analyzed with regard to the question whether it is generally appropriate and appealing for human users or whether individual users prefer particular nonverbal behaviours. Also, it can be asked whether people vary with regard to the desired number of certain behaviours (e.g., the agent´s smiling or gaze). Inter-individual differences with regard to the perception of nonverbal behaviour seem plausible against the background of, for example, the increased sensitivity of women with regard to nonverbal behaviour [5]. Thus, the goal of the present paper is to

J. Allbeck et al. (Eds.): IVA 2010, LNAI 6356, pp. 468–474, 2010.

analyze whether different user groups exhibit specific preferences for specific nonverbal behaviours of an interface agent. We present evidence from three different studies which originally targeted the question which effects different nonverbal behaviours of the agent Max [6] have: Eyebrow raising and gestural activity [7], smiling [8] and feedback behaviour [9]. As we additionally assessed information on the user (gender, age, and computer literacy) we are now able to present an overview on the influence of these variables on the evaluation of specific nonverbal cues and can derive first conclusions on the preferences of different user groups with regard to the nonverbal design of embodied agents.

2 Method

Three studies. Three studies were conducted, employing the ECA Max as stimulus material, to test whether an agents´ nonverbal behavior has different effects on different users. The experimental sessions were conducted in the laboratory at Bielefeld University (Germany) where Max [6] is displayed on a life-sized screen.

Procedure. The procedure was similar in all three studies: Each participant had a 5-10 minute lasting small-talk with Max, which was recorded via video camera. Since speech technology is so far not able to reliably recognize spoken language input in noisy environments, a so-called "Wizard of Oz" scenario was used. The Wizard, however, had no influence on the reactions of Max, which were autonomously generated by the system. After the conversation participants had to fill in a questionnaire concerning the feelings they had during the interaction, their perception of Max, and their evaluation of the interaction in general. The participants` gender, age and computer literacy were assessed as well. With regard to age, however, merely study 1 tested a relevant number of participants older than 30 years and therefore age was only considered as potentially moderating variable in this study.

Dependent Variables. The dependent variables were collected via paper-and-pencil questionnaires and via video-recordings of the participants during the interaction with Max. The questionnaires were designed based on existing questionnaires, which were developed and repeatedly employed in previous studies [7]. By means of 20-items, the feelings of the participants during the interaction were determined (e.g. "attentive", "amused", "relaxed" and "lethargic"). Participants had to state their level of agreement by means of 5-point Likert-scales, with the extremes "strongly disagree" and "strongly agree". The person perception of Max was measured by means of a 34-item semantic differential (seven-point bipolar rating scales, whose extremes are designated by two opposite adjectives). It was assessed, for example, whether Max was perceived rather as "warmhearted" or "cold", "self-confident" or "shy", "wooden" or "animated". Finally, the general evaluation of the interaction was assessed by items which asked for the participants` enjoyment of the interaction, the perceived controllability of the conversation, and whether participants could imagine to complete tasks like programming a video recorder with Max. Again, participants had to state their level of agreement by means of 5-point Likert-scales. The participants` nonverbal behavior towards the agent was assessed by quantitative analyses of the recorded video material.

Independent Variables: Manipulations of the agent´s nonverbal behavior

Study 1: Eyebrow movement & self-touching gestures. 50 persons, 28 female and 22 male, aged from 15 to 72 years volunteered to participate in the study (mean value = 27.73 years; SD = 11.53). They were confronted with Max, who´s nonverbal behavior was manipulated with respect to eyebrow movements and self-touching gestures. The presence of eyebrow movement as well as self-touching gestures were manipulated [7].

Study 2: Frequency of smiling. 104 persons (52 male, 52 female) aged from 19 to 55 years participated in the second study. In order to analyze whether the frequency of an agent´s smile has an effect on its evaluation, three conditions were distinguished: a) no-smile condition, in which Max did not show any smile; b) infrequent smile condition, in which Max shows occasional smiles, and c) frequent smile condition, where Max shows frequent smiles [8].

Study 3: Different styles of feedback. In the third study 70 persons, 19 males and 51 females, aged from 17 to 48 years (M= 24.09, SD= 5.717) participated. In this study different styles of the agent´s feedback were varied: emotional feedback, which provided a feedback about the emotional state of Max (including smiles and compliments), and envelope feedback, which provided a feedback about the comprehension of the participants´ contributions and presents Max as an attentive listener [9].

3 Results

3.1 Feelings during the Interaction with Max

Effects moderated by participants` gender

Study 1: When we added the participants` gender into an ANOVA, no main effects, but a significant interaction was found for the factor „interest". Females show more interest than men, when self-touching gestures were shown ($F(1;42)= 6.272$; $p = .016$; Part. $\eta^2 = .130$; female: M = 0.56; SD = 0.77; male: M = -0.55; SD = 1.21). In turn when no self-touching gestures were shown, male participants experienced the conversation as more interesting than females (male: M= 0.23, SD= 1.00; female: M= -0.16, SD= 0.83).

Study 2: In study 2 one main effect of gender emerged ($F(1; 103)= 5.67$; $p= .019$; Part. $\eta^2= .060$), showing that women (M= -0.29; SD= 0.85) are more nervous than men (M= 0.21; SD= 1.12) when interacting with Max. There was also one interaction between gender and Max`s smiling concerning the factor disinterest ($F(2; 103)= 4.72$; $p= .011$; Part. $\eta^2 =.10$). Whereas men are least disinterested in the no-smile condition (M=- 0.41; SD= 0.50), women are most disinterested in this condition (M= 0.06; SD= 1.04). In the infrequent-smile condition, the pattern is reversed (males: M= 0.57; SD= 0.98; females: M= -0.48; SD= 0.89). However, women and men do not differ in disinterest in the frequent-smile condition (males: M= 0.05; SD= 1.31; females: M= -0.05; SD= 0.74).

Study 3: No significant differences or interactions occurred due to gender.

Effects moderated by participants` age

Study 1: As depicted above, age only varied considerably in study 1, therefore it was only considered within this study. Here, the participants` age had an effect on their feelings during the interaction, explicitly on the factor nervousness, $F(1; 40)= 4.253$; $p= .046$; Part. $\eta^2 = .096$. Younger participants (15-23 years) were significantly less nervous (M= -0.34, SD= 0.69) than older participants (24-72 years) (M= 0.288, SD= 1.108). Additionally, study 1 yielded a significant interaction between age and self-touching gestures for the factor attention, $F(1; 40)= 5.019$; $p= .031$; Part. $\eta^2 = .111$: Younger participants were more attentive (M= 0.438, SD= 0.600) than older participants (M= -0.448, SD= 1.448), when Max showed self-touching gestures. In turn, older participants were more attentive (M= 0.31, SD= 0.77) than younger ones (M= -0.04, SD= 0.41), when no self-touching gestures were shown.

Effects moderated by participants` computer literacy

Study 1: We observed that experts were less nervous (M= -0.20; SD= 0.89) during the interaction with Max than novices (M= 0.33; SD= 1.11). This was proven by a significant main effect for the factor nervousness ($F(1; 42)= 6.736$; $p= .040$; Part. $\eta^2 = .097$).

Study 2 and 3: Neither main effects nor interaction effects of the users` computer literacy emerged.

3.2 Evaluation of Max

Effects moderated by participants` gender

Study 1 and study 3: The gender of the participants had no significant influence on the perception of Max nor interacted with the independent variables.

Study 2: Results for the factor incompetence show a significant main effect, ($F(1; 103)= 5.761$; $p= .018$; Part. $\eta^2= .061$): Max`s incompetence is rated higher by male participants (M= 0.29; SD= 1.02) than by female participants (M= -0.23; SD= 5.76). Thus, men are more rigorous in their judgment with regard to competence. There was no significant interaction between gender and the smiling condition.

Effects moderated by participants` age. In study 1, no main effects, or interactions were found.

Effects moderated by participants` computer literacy

Study 1: With regard to computer literacy, experts experienced Max as more strained (M= 0.20, SD= 0.86) than novices (M= -0.32, SD= 1.16), as a main effect for the factor strain showed, $F(1; 42)= 4.235$; $p= .046$; Part. $\eta^2 = .092$.

Study 2: Also study 2 emphasizes differences in the evaluation of Max due to computer literacy: the factor passiveness was significantly ($F(2; 103)= 4.219$; $p= .018$; Part. $\eta^2 = .087$) stronger perceived by gamers (M= 0.38, SD= 1.11) than by experts (M= 0.26, SD= 1.01) and least by novices (M= -0.29; SD= 0.86). No significant interactions between Max`s smiling behavior and computer literacy were found.

Study 3: No main effects, or interactions were found.

3.3 General Evaluation of the Interaction

Effects moderated by participants` gender

Study 1: ANOVA yielded a significant interaction between self-touching gestures and the participants` gender for the factor acceptance: When self-touching gestures were shown by Max women show a higher acceptance than men ($F(1; 42)= 4.552$; $p= .039$; Part. eta 2= .098; woman: M= 0.57, SD= 0.99; men: M= -0.38, SD= 1.10). But when no self-touching gestures were shown men show more acceptance than women (men: M= 0.13, SD= 0.94; woman: M= -0.21, SD= 0.87).

Study 2 and 3: No main effects or interactions on factor level were observable.

Effects moderated by participants` age. Neither main effects, nor interaction effects were found in study 1.

Effects moderated by participants` computer literacy. No main effects, or interactions emerged on factor level in any of our studies.

3.4 Participants` Nonverbal Behavior towards the Agent

Effects moderated by participants` gender

Study 1: No main effects but a significant interaction for the participants` gender were found: Self-touching gestures by the participants depend on eyebrow raising of the agent, as study 1 shows. But the direction differs between male and female participants. The interaction effect ($F(1; 42)= 4.095$; $p= .049$; Part. $\eta^2 = .089$) suggests that female participants show more self-touching gestures (M= 0.12, SD= 0.17) than male participants (M= 0.03, SD= 0.05), when Max did not raise his eyebrows. Instead, females show less self-touching gestures (M= 0.09, SD= 0.10) as men (M= 0.13, SD= 0.11) when Max shows eyebrow raising behavior.

Study 2: One main effect of the participants` gender was found in study 2, but no significant interaction emerged: women showed more ($F(1; 103)= 9.8$; $p = .002$; Part. $\eta^2 = .087$; men: M= 0.49; SD= 0.319; women: M= 0.70; SD = 0.34) and longer ($F(1; 103)= 6,34$; $p= .011$; Part. $\eta^2= .061$; men: M= 0.07, SD= 0.06 ; women: M= 0.11; SD = 0.07) full smiles than men.

Study 3: No main effects or interactions emerged.

Effects moderated by participants` age. In study 1 the age of the participant caused no main, or interaction effects.

Effects moderated by participants` computer literacy

Study 1 and study 3: No significant main effects or interactions were observable.

Study 2: An ANOVA with respect to the influence of the participants` computer literacy on their nonverbal behavior showed one main effect, but no interaction. The main effect occurred for the overall duration of smiling behaviors of the participants: novices smiled the most followed by experts and gamers ($F(2; 103)= 3.333$; $p= .040$; Part. $\eta^2 = .062$; novices: M= 0.44; SD= 0.17; experts: M= 0.38; SD= 0.16); gamers: M= 0.34; SD= 0.15).

4 Discussion

The aim of this paper was to analyze whether different user groups exhibit specific preferences for specific nonverbal behaviors of an interface agent. For that purpose, evidence from three empirical studies with the agent Max was presented, with a focus on moderating aspects, namely the participants` gender, age and computer literacy. The results show that the attributes have an influence on the feelings of the participants during their interaction with Max, on the evaluation of Max, on the general evaluation of the interaction, as well as on the participants` nonverbal behavior. In the following, we sum up and interpret the corresponding results in detail.

Know your users` gender. Men and women have different preferences with regard to embodied agents. In fact, compared to the effects of age and computer literacy, the influence of gender was prevailing. When developing agents it would thus be most important to take these results into account – either by providing two versions from which the user can choose, or by designing the agent accordingly when it is known whether predominantly men or women will use it.

Study 2 showed that women were, in general, more nervous during the interaction with the agent, which is on the one hand not astonishing given that women are still less accomplished with regard to computer technology, but on the other hand it contradicts the vision that embodied agents will facilitate human-computer-interaction for these kinds of users. Our data suggests that female users´ interest and acceptance can be increased when self-touching gestures are implemented (see study 1) and when the agent frequently smiles. Therefore we conclude that women especially benefit from an increased nonverbal behavior of the agent, in line with the finding that women are more sensitive for nonverbal behaviors [5]. One can speculate that they are reassured when they perceive signals from the interlocutor that they can interpret.

Besides, for male users it was shown that self-touching gestures decrease interest in the conversation with an agent (see study 1). Also men in study 2 were most interested when the agent did not smile. Thus it can be concluded that agents for male users should be kept simple with regard to nonverbal behavior.

Know your users` age. Study 1 showed that also older persons were more nervous when they interacted with Max. Additionally the interaction between the age of the participants in study 1 and the manipulation of self-touching gestures demonstrates that older participants were more attentive when no self-touching gestures were shown. Especially the increased nervousness should be analyzed carefully in future studies in order to be able to derive conclusions.

Know your users` computer literacy. Computer novices proved to be more nervous than other users. This is in line with previous findings that computer laypeople do not benefit from embodied agents in the way that it is typically hoped for [10]. Also, people who are frequently in contact with computers and newer technologies have different demands concerning embodied agents than persons who are novices in this area. The results of our studies showed that computer experts evaluate Max as more strained and gamers as more passive. This might be due to the fact that especially gamers are used to sophisticated virtual characters. As this is independent from the specific nonverbal behavior no specific design guidelines for nonverbal behavior can

be derived but still the results point to the importance of supporting the specific needs of novices versus experts.

Conclusions. Although it has to be considered that results could have been more consistent since not all studies yielded the same main effects, they show distinct patterns. The results presented here are just a first step towards knowing the user but they already indicate that there are differences which when taken into account might lead to specific design guidelines. As suggested above this might either lead to different versions of an embodied agent from which the users can choose or – if the agent is to be employed with only one specific group of users – to a specifically designed version of the agent. As especially gender effects were prevailing these might be a good starting point for first implementations and further tests.

Besides, other moderating aspects like culture should be taken into account in future studies, because especially different cultures might like to see different nonverbal behaviors, as also people from different cultures show and prefer different behaviors.

References

1. Shneiderman, B.: Designing the user interface: Strategies for effective human-computer interaction, 1st edn. Addison-Wesley, Reading (1987)
2. Cassell, J., Bickmore, T., Campbell, L., Vilhjálmsson, H., Yan, H.: Human conversation as a system framework: Designing embodied conversational agents. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (eds.) Embodied conversational agents, pp. 2–63. MIT Press, Cambridge (2000)
3. Krämer, N.C.: Social communicative effects of a virtual program guide. In: Panayiotopoulos, T., Gratch, J., Aylett, R.S., Ballin, D., Olivier, P., Rist, T. (eds.) IVA 2005. LNCS (LNAI), vol. 3661, pp. 442–453. Springer, Heidelberg (2005)
4. Domagk, S.: Pädagogische Agenten in multimedialen Lernumgebungen. Empirische Studien zum Einfluss der Sympathie auf Motivation und Lernerfolg. Logos, Berlin (2008)
5. Hall, J.A.: Nonverbal sex differences. Communication accuracy and expressive style. Johns Hopkins University Press, Baltimore (1984)
6. Kopp, S., Gesellensetter, L., Krämer, N.C., Wachsmuth, I.: A conversational agent as museum guide - design and evaluation of a real-world application. In: Panayiotopoulos, T., Gratch, J., Aylett, R.S., Ballin, D., Olivier, P., Rist, T., et al. (eds.) IVA 2005. LNCS (LNAI), vol. 3661, pp. 329–343. Springer, Heidelberg (2005)
7. Krämer, N.C., Simons, N., Kopp, S.: The effects of an embodied conversational agent's nonverbal behavior on user's evaluation and behavioral mimicry. In: Pelachaud, C., Martin, J.-C., André, E., Chollet, G., Karpouzis, K., Pelé, D. (eds.) IVA 2007. LNCS (LNAI), vol. 4722, pp. 238–251. Springer, Heidelberg (2007)
8. Krämer, N.C., Kopp, S., Sommer, N., Becker-Asano, C.: Smile and the world will smile with you. The effects of a virtual agent's smile on users' evaluation and behavioral mimicry (revised and resubmitted)
9. Von der Pütten, A., Reipen, C., Wiedmann, A., Kopp, S., Krämer, N.C.: Comparing emotional vs. envelope feedback for ECAs. In: Prendinger, H., Lester, J.C., Ishizuka, M. (eds.) IVA 2008. LNCS (LNAI), vol. 5208, pp. 550–551. Springer, Heidelberg (2008)
10. Krämer, N.C., Rüggenberg, S., Meyer zu Kniendorf, C., Bente, G.: Schnittstelle für alle? Möglichkeiten zur Anpassung anthropomorpher Interface Agenten an verschiedene Nutzergruppen. In: Herzceg, M., Prinz, W., Oberquelle, H. (eds.) Mensch und Computer 2002, pp. 125–134. Teubner, Stuttgart (2002)