

A second chance to make a first impression? How appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time

Kirsten Bergmann^{1,2}, Friederike Eyssel¹, and Stefan Kopp^{1,2}

¹ Center of Excellence in “Cognitive Interaction Technology” (CITEC), Bielefeld University

² Collaborative Research Center 673 “Alignment in Communication”, Bielefeld University

P.O. Box 10 01 31, 33501 Bielefeld, Germany

{kbergman, skopp}@techfak.uni-bielefeld.de

friederike.eyssel@uni-bielefeld.de

Abstract. First impressions of others are fundamental for the further development of a relationship and are thus of major importance for the design of virtual agents, too. We addressed the question whether there is a second chance for first impressions with regard to the major dimensions of social cognition—warmth and competence. We employed a novel experimental set-up that combined agent appearance (robot-like vs. human-like) and agent behavior (gestures present vs. absent) of virtual agents as between-subject factors with a repeated measures design. Results indicate that ratings of warmth depend on interaction effects of time and agent appearance, while evaluations of competence seem to depend on the interaction of time and nonverbal behavior. Implications of these results for basic and applied research on intelligent virtual agents will be discussed.

Keywords: Evaluation, agent appearance, nonverbal behavior, warmth, competence

1 Introduction

One of the major challenges in current research on intelligent virtual agents (IVAs) is the question of how to elicit positive affect, user acceptance and even long-term relationships between users and agents. Inspired by human-human interaction it seems to be of striking importance that a user’s *first impression* of an IVA is positive: When two people meet for the first time, they immediately form initial ideas from each other. These early impressions have a major impact on how their relation further develops. People’s behavior towards others is shaped depending on differences in first impressions such that people who have favorable impressions of someone tend to interact more with that person than others having unfavorable impressions [18]. First impressions are, therefore, an important basis for whether humans will build rich relations with others.

The tendency to form first impressions is absolutely fundamental with regard to person perception and social cognition [28]. Thus, the two universal dimensions of social cognition—*warmth* and *competence* [12, 10]—are notably shaped by the first moments of contact, whereby the warmth dimension captures whether people are friendly and

well-intentioned and the competence dimension captures whether people have the ability to deliver on those intentions. A number of studies have shown that warmth and competence assessments determine whether and how we intend to interact with others (cf. [10]): We seek the company of people who are assumed to be warm and avoid those who appear less sociable (i.e. cold). With regard to competence, we prefer to cooperate with people we judge as competent, while incompetent people are disregarded. With regard to the relation of warmth and competence, Fiske et al. [12, p. 77] state that “warmth judgments are primary: warmth is judged before competence, and warmth judgments carry more weight in affective and behavioral reactions”. In addition, “perceived warmth is more easily lost and harder to regain compared to perceived competence” [10, p. 17].

So, as early evaluations of others are a major concern for social evaluations—which cues do people take into account when making first impressions? How can we strengthen the chances for making a solid first impression? Empirical evidence from social psychology has demonstrated that initial impressions are formed rather quickly on the basis of minimal information with visual *appearance* and *nonverbal behavior* providing the major cues [25]. With regard to the latter, *co-speech gestures* are a particular kind of nonverbal behavior: Gestures convey semantic information, while the main function of other nonverbal behaviors like facial expressions or body posture is to communicate sympathy/antipathy and other affective/emotional signals. Accordingly, gestures are especially important in settings and tasks in which information is to be transferred and we will focus on this kind of nonverbal behavior here. Nevertheless, despite of all evidence for the fact that first impressions are lasting impressions in terms of setting the tone for a relationship, they do not define its boundaries or potential. Subsequent encounters still have the chance to modify real quality of any relationship. Especially competence traits are still subject to subsequent modifications (cf. [10]). So it seems that there is still a chance for what we call a ‘second impression’.

All these issues of first and second impressions are, however, mostly unexplored for virtual agents. Although there is empirical evidence from IVA research that provides support for the fact that visual appearance and nonverbal behavior are cues of major importance for human’s evaluation of virtual agents (cf. [21]), the relation of (1) agent-related cues like agent appearance and nonverbal behavior, (2) social evaluation in terms of the major dimensions of warmth and competence as well as (3) dynamic modifications of impressions has not been investigated for IVAs, yet. In this paper, we aim to investigate interaction effects of these variables. In particular, we address the question how warmth and competence ratings change from a first impression after a few seconds to a second impression after a longer period of human-agent interaction depending on manipulations of agents’ visual appearance and nonverbal behavior. The following section gives an overview of related work and background literature. Section 3 describes the setting and procedure of the evaluation study. Results are presented in Section 4. Finally, we discuss the results and draw conclusions in Section 5.

2 Related Work and Background

2.1 Effects of agent appearance

A growing body of empirical evidence suggests that the appearance of virtual and robotic agents has an extensive influence on how humans evaluate them, as Goetz et al. [14, p. 1] put it aptly: “The book is judged by its cover. [...] A robot’s appearance and behavior provide cues that influence perceptions of the robot’s propensities, and assumptions about its capabilities”. A major dimension of agent appearance which is subject to a large amount of recent research activities (particularly in the field of robotics) is the degree of *human-likeness*. Among researchers it is a controversial issue whether machines should be endowed with a human-like interface or not. On the one hand, it is argued that humanoids provide a more intuitive interface because rules of human interaction can easily be transferred [7]. On the other hand there are also opponents who argue that a human-like appearance results in unrealistic expectations or even fear [11]. In a few studies from robotics it was examined how human-like vs. machine- or robot-like agents are evaluated by human users, but none of them directly investigated the variables of warmth and competence. Hinds et al. [16] found that machine-like interfaces tend to be treated less politely and less socially interactive than human-like interfaces in a joint task between humans and robots. Moreover, expectations regarding abilities and reliability were lower for the machine-like interfaces. Nishio & Ishiguro [24] found a strong effect of appearance on human evaluations: They reported that interlocutors tend to hold different impressions of robotic agents of different appearances, even if the agents were tele-operated by a single person. Goetz et al. [14] presented evidence for a task-dependent relation between a robot’s appearance and users’ acceptance of and cooperation with a robot. In fact, participants systematically preferred robots for jobs when the robot’s human-likeness matched the sociability and seriousness required in those tasks. Woods et al. [30] further provided evidence that children judge human-like robots as aggressive, but machine-like robots as friendly.

Often discussed with regard to effects of appearance and behavior is Mori’s ‘uncanny valley’ hypothesis [22] which states that the perceived familiarity of robots rises with increasing anthropomorphism until a point is reached beyond which ratings go into reverse and robots are perceived as eerie instead of familiar. According to Mori the uncanny valley effect is even stronger for animated and moving agents than it is for static agents. Although originally proposed for robotics this hypothesis has also been transferred to virtual agents where a couple of studies provided evidence for the effect in IVAs as well [13, 2]. It turned out that the degree of realism does not necessarily result in positive evaluations. Instead, it is more important that the degree of realism is consistent with the agent’s behavior.

In another line of research it is investigated how virtual agents are perceived depending on manipulations of human-likeness in contrast to zoomorphic agents. Sträfling et al. [29] compared a cartoon-like rabbit and a realistic anthropomorphic agent in a teaching scenario. Results showed that the appearance of the agent mattered such that the rabbit-like agent was preferred. By contrast, Bailenson et al. [1] did not find significant differences between a human-like and a rabbit-like appearance with regard to perceived likeability, but both interfaces received significantly higher ratings than a third

abstract IVA variant. In a different study Buisine & Martin [8] compared three different cartoon-like agents (two male, one female) and reported evidence for an influence of agents' appearance on likeability: one of the male agents was significantly preferred over the others. It was, however, unclear by which aspect(s) of appearance this effect was caused. An analysis of participants' qualitative comments suggests that this effect was due to the wide smile of that particular character.

Overall, results of studies investigating how virtual agents are perceived depending on their visual appearance are inconclusive (see [21] for a more comprehensive overview) which may be due to the fact that appearance consists of many different aspects and cannot be broken down to a single variable. In addition, as some researchers pointed out, agent appearance should be consistent with the agent's behavior [13, 2] and also with the task [14]. Accordingly, it makes sense to investigate the effect of agent appearance in interaction with other variables. Nevertheless, the conclusion can be drawn that agent appearance has the potential to modify ratings of warmth, likeability etc.

2.2 Effects of nonverbal behavior

Research investigating the impact of IVA's nonverbal behavior on human perception of warmth and competence is only sparse. So far, only Niewiadomski et al. [23] considered the role of verbal vs. nonverbal vs. multimodal emotional displays on warmth, competence and believability, whereby nonverbal behavior consisted of facial expressions accompanied by emotional gestures. It was found that all dependent measures (warmth, competence, and believability) increased with the number of modalities used by the agent. From other studies, although often not directly addressing the dimensions of warmth and competence, there is further evidence that endowing virtual agents with human-like, nonverbal behavior may lead to enhancements of the likeability of the agent, trust in the agent, satisfaction with the interaction, naturalness of interaction, ease of use, and efficiency of task completion [6, 15].

With regard to the particular effects of co-speech gestures, Krämer and colleagues [20] found no effect on agent perception when comparing a gesturing agent with a non-gesturing one. The agent displaying gestures was perceived just as likeable, competent, and relaxed as the agent that did not produce gestures. In contrast, a number of other studies found beneficial effects of virtual agents using gestures. Cassell & Thorisson [9] reported that nonverbal behavior (including beat gestures) resulted in an increase of perceived language ability and life-likeness of the agent, as well as smoothness of interaction. A study by Rehm & André [26] investigated whether a gesturing agent would change the perceived politeness tone compared to that of the textual utterances and whether the subjective rating is influenced by the type of gestures (abstract vs. concrete). These studies revealed that the perception of politeness depends on the graphical quality of the employed gestures. In cases where iconic gestures were rated as being of higher quality than metaphoric gestures, a positive effect on the perception of the agent's willingness to co-operate was observed. In cases where iconic gestures were rated as being of lower quality than metaphoric gestures, a negative effect on the perception of the agent's willingness to co-operate was observed. Moreover, in the aforementioned study by Buisine & Martin [8], effects of different types of speech-gesture cooperation in an agent's behavior were found: redundant gestures increased ratings of explanation

quality, expressiveness of the agent, likeability and positive perception of the agent's personality. Finally, in our own previous research we compared different kinds of gestural behavior in an IVA (individualized gesture models vs. common gesture models vs. control conditions) [5]. Especially gestural behavior generated with individualized models turned out to have significant benefits on human ratings in terms of likeability, competence and human-likeness.

In sum, previous research has shown that IVAs using gestures provide the potential to increase human judgements of likeability, competence and other variables. However, gesture use alone is not a guarantee for positive ratings. These are rather depending on further factors like gesture quality [26] or adequate gesture models.

2.3 Changing Impressions

Regarding the question whether user evaluations of robotic or virtual agents might change as a function over time research is very sparse. There is only work by Komatsu & Yamada [19] who investigated what they called 'adaptation gap'—the difference between users' expectations before starting their interactions with a robot and their evaluation after interacting. Participants were divided into two groups which received different information about the interface beforehand (lower expectation group vs. higher expectation group). The study revealed a significant difference between experimental conditions with regard to participants' post-interaction judgements.

Thus, there is initial evidence that human users' judgements of agents might modify with ongoing interaction. There is, however, no detailed account of interaction effects between dynamics of evaluation and other variables such as agents' visual appearance and behavior. Our study presented in this paper aims to be a first step into the direction of closing this gap.

3 Study

The study employed a 2 x 2 between-subject factorial design with repeated measures to investigate the effects of agent appearance (robot-like vs. human-like) and agent behavior (gestures absent vs. gestures present) on the perception of virtual agents by human participants.

Participants A total of 80 participants (20 in each condition), aged from 19 to 48 years ($M = 25.11$, $SD = 5.45$), took part in the study. 51 participants were female and 29 were male. All of them were recruited at Bielefeld University and received 3 Euros for participating.

Procedure The experiment was conducted in two consecutive phases. In the first phase, participants were provided with a short introduction by the virtual agent which took approximately 15 seconds: "Hello, my name is Billie/Vince. In a moment you will have the have the possibility to getting to know me closer. But first you will be provided with a questionnaire concerning your first impression of myself". Subsequently, participants were asked to state their first impression regarding their perception of the virtual

agent’s personality in a questionnaire ($T1$, took approximately five minutes). In the second phase, the virtual agent described a building with six sentences which took approximately 45 seconds. Each sentence was followed by a pause of three seconds. Participants were instructed to carefully watch the presentation given by the virtual agent in order to be able to answer questions regarding content and subjective evaluation of the presentation afterwards. Immediately upon receiving the descriptions by the IVA, participants filled out a second questionnaire ($T2$) stating their perception of the virtual agent’s personality at this point of time.

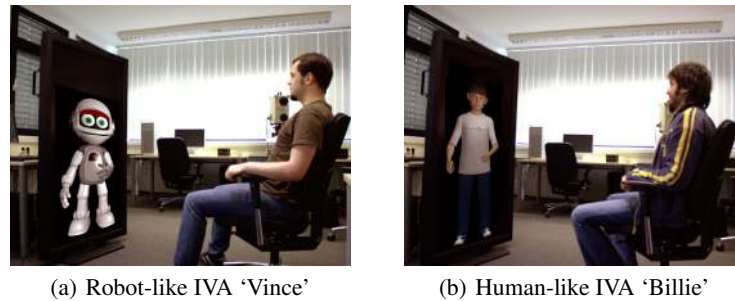


Fig. 1. Set-up of the study in which the IVAs provided users with explanations about a building.

Figure 1 shows the setup used for stimulus presentation: The virtual agent was displayed on a 80 x 143 cm screen. Participants were seated 170 cm away from the screen and their heads were approximately leveled with the virtual agent’s head. Participants have been left alone for the stimulus presentations, and after receiving the questionnaires to complete them, i.e., neither the experimenter nor the virtual agent were present.

Independent Variables As a within-subject variable repeated measurements were taken at two *points of measurement* ($T1$, $T2$) in the experimental procedure. $T1$ was chosen to measure how users rate the agent in terms of a first impression after 15 seconds, while $T2$ was chosen to evaluate user ratings after a longer time of information presentation by the IVA (second impression). In addition, participants were randomly assigned to one out of four conditions which resulted from the manipulation of two independent between-subject variables: *agent appearance* and *agent behavior* (gesture use).

Agent appearance Two different kinds of virtual agents were employed in the study with regard to the dimension of human-likeness (anthropomorphic–robotic IVAs). We employed the IVA ‘Vince’ as a robot-like character and ‘Billie’ as a human-like agent (see Figure 1). To hold conditions as constant as possible despite the intended manipulations, both agents were displayed with the same overall size, and no other non-verbal behaviors than gesture use (see below) were employed. Verbal utterances were also identical across conditions, although the two agents used different synthetic voices (both Mary TTS [27]) each of which matched with the agent’s visual appearance. A

child-like voice was used for Billie which is based on a female German Hidden semi-Markov model voice. It was modified by a slightly shortened vocal tract, F0 values were shifted by 50 Hz resulting in a higher-pitched voice, the range of f0 values was expanded and duration scaling was modified such that the synthesized speech output was slower (children speaking slower than adults). A machine-like voice for Vince was based on a male German Hidden semi-Markov model voice. This voice was also modified by a slightly shortened vocal tract, F0 values were shifted by 100 Hz and the range of f0 values was expanded.

Agent behavior The gesturing behavior of the virtual agents was manipulated in the way that either no gestures were used at all ('gestures absent'), or the *Generation Network for Iconic Gestures* (GNetIc, for details, see [3]) was employed to generate gestures for the virtual agents ('gestures present'). The GNetIc model serves to simulate co-speech gesture use in the style of individual speakers accounting for obvious and important inter-individual differences in gestural behavior. For the purpose of the current study, we employed an individual speaker's GNetIc model which has been shown to increase the perceived quality of an object description given by a virtual human, and also resulted in a more positive rating of an IVA in terms of likeability, competence and human-likeness, compared to other GNetIc models as well as control conditions [5]. All descriptions given by the virtual agents were produced fully autonomous at runtime by using an integrated speech and gesture production architecture [4].

Dependent Variables Participants' responses regarding the dependent variables were collected using 18 items [12, 17] such as 'pleasant', 'friendly', 'helpful' which had to be assessed on seven-point Likert scales how well they apply to Billie/Vince (not appropriate–very appropriate). For subsequent data analyses, average scores were computed as indices for the scales of warmth and competence, whereby remaining ones were used as filler items.

WARMTH The following eight items measured warmth-related traits: pleasant, sensitive, friendly, helpful, affable, likeable, approachable, sociable. The scale was highly reliable with Cronbach's $\alpha = .92$ for T1 and $\alpha = .92$ for T2.







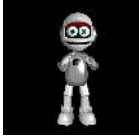









COMPETENCE Competence-related traits were measured with four items (intelligent, organized, expert, thorough). The competence scale also showed high reliabilities of Cronbach's $\alpha = .84$ for T1 and $\alpha = .83$ for T2.

Data Analysis First, descriptive data analyses were conducted to summarize the data. Second, a mixed-model design ANOVA was conducted to analyze the effect of the within-subject variable *point of measurement* (T1 vs. T2) and between-subject variables *agent appearance* (robot-like vs. human-like) and *agent behavior* (gestures absent vs. gestures present) on the dependent variables WARMTH and COMPETENCE.

4 Results

The descriptive data analysis revealed means and standard deviations as summarized in Table 2. Concerning WARMTH, the robot-like agent with gestures present had the

Table 1. Stimuli presented in the ‘gestures present’ conditions: verbal description given in each condition (left column; translated to English; gesture positions labelled with squared brackets) and the different virtual agents either displaying gestural behavior.

Agent appearance	<i>robot-like</i>	<i>human-like</i>
Introduction		
[The church is squared]...		
...and in the middle there is [a small spire.]		
The spire has [a tapered roof].		
And the spire has [a clock].		
There is [a door] in front.		
And in front of the church there is [a low, green hedge].		
There is [a large deciduous tree] to the right of the church.		

highest mean score for both points of measurement. However, lowest WARMTH-ratings at measurement point T2 were observed for the robot-like agent as well, namely in the condition with gestures absent. For T1 the human-like agent with gestures received lowest WARMTH-ratings. With regard to competence the highest mean scores at both points of measurement were reached by the robot-like agent with gestures present as well as the human-like agent with gestures absent.

In the following we report (interaction) effects of between-subject and within-subject variables with regard to the two dependent measures of WARMTH and COMPETENCE. Only significant effects are reported at a level of $p < .05$.

Table 2. Means and standard deviations of WARMTH and COMPETENCE as a function of *agent appearance*, *agent behavior* and *point of measurement*.

		Robot-like virtual agent		Human-like virtual agent	
		Gestures absent	Gestures present	Gestures absent	Gestures present
		M (SD)	M (SD)	M (SD)	M (SD)
Warmth	T1	4.56 (1.09)	5.15 (1.21)	4.25 (1.14)	4.05 (1.18)
	T2	3.90 (1.32)	4.49 (1.36)	4.12 (1.09)	4.19 (1.18)
Competence	T1	4.50 (0.93)	4.75 (1.35)	4.85 (1.48)	3.98 (1.21)
	T2	4.38 (1.26)	4.93 (1.16)	4.79 (1.42)	4.55 (1.11)

WARMTH There was a significant main effect of the within-subject variable *point of measurement* on the degree of WARMTH perceived by participants ($F_{(1,76)} = 11.65, p = .001$) in the way that at measuring point T1 the agents were perceived as warmer than at measuring point T2.

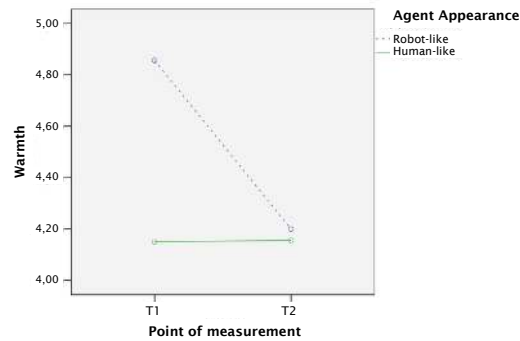


Fig. 2. WARMTH ratings as a function of *agent appearance* and *point of measurement*.

Concerning interaction of independent variables, there was a significant interaction effect of *point of measurement* and *agent appearance* ($F_{(1,76)} = 12.10, p = .001$) indicating that the rating of WARMTH at the two points of measurement differed significantly. As visualized in Figure 3 ratings of WARMTH decreased for the robot-like agent between the two points of measurement, while ratings remained constant for the human-like agent. While participants rated the robot-like agent as being warmer than the human-like agent at measuring point T1, ratings of warmth did not differ between robot-like and human-like agents at T2.

COMPETENCE Regarding perceived COMPETENCE of the virtual agents, there was a significant effect for the interaction of the *point of measurement* and *agent behavior* ($F_{(1,76)} = 4.56, p = .04$). While gesture use was found to result in an increase of perceived competence of the virtual agents between the two points of measurement, ratings of agent competence slightly decreased when the agents did not use any gestures (see Figure 3).

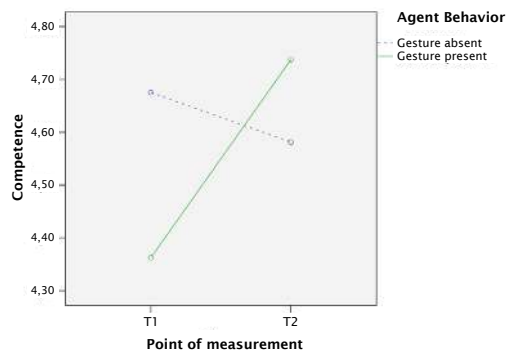


Fig. 3. COMPETENCE ratings as a function of *agent behavior* and *point of measurement*.

5 Conclusion

The goal of this paper was to explore first and second impressions human users form of virtual agents depending on agent appearance and nonverbal (gestural) behavior. To investigate this objective, we conducted an experiment with a mixed design. Our results speak to three important issues. First of all, participants' rating with regard to the agents' warmth was sensitive to the point of measurement for the robot-like agent. According to participants' first impression the agents were perceived as warmer than according to second impressions after a longer phase of information presentation by the IVAs. By contrast, there were no significant effects of the point of measurement observed for competence. Second, with regard to warmth, there was a significant interaction effect of

the point of measurement and agent appearance. While warmth-ratings between the two points of measurement decreased significantly for the robot-like agent, ratings remained constant for the human-like agent. Third, regarding competence, we found a significant interaction effect of the point of measurement and agent behavior. That is, the way that gesture use helped to increase competence ratings between measurement points, while the absence of gestures resulted in a decrease of competence ratings.

Compared to related work from robotics and virtual agents, our results provide further support for the fact that human users' evaluation are actually sensitive regarding the agents' appearance and nonverbal agent behavior. Specifically, our findings are generally in line with previous results from IVA research by Niewiadomski et al. [23] who showed that ratings of agents' warmth and competence increase with the number of modalities used by the agent. The particular setup of our study, however, allows for a more sophisticated view showing that the two dimensions of social cognition are not dependent on the agents' behavior (and agent appearance) in equal measures. Furthermore, going beyond previous work, we showed that IVA ratings are sensitive to the point of measurement.

How can we explain these findings? The overall decrease of warmth and increase of competence (depending on presence/absence of co-speech gestures) is in accordance with evidence from social psychology ([12, 10], see Sect. 1), stating that warmth judgements are made rather quickly and that warmth is easily lost and hard to regain in contrast to competence. However, the interaction effect of the point of measurement and agent appearance is notable. Why do warmth rating decrease only for the robot-like agent? We suppose that this is a matter of expectations. The robot-like agent Vince with his large head and big eyes fulfills quite some characteristics of a child-like look which is typically associated with many positive attributes. In addition, in the graphical representation of Vince appears to be much closer and present than Billie. It seems, however, that being exposed to this interface for a longer time results in a decrease of this effect. We suggest, that the particular setting we employed in this study might have had an impact, too. In a different task, e.g. enforcing more interaction or personal communication, warmth ratings are likely to be different than in our setting of information providing by the IVA.

What can we learn from these findings for the design of (interactions with) virtual agents? First of all, our results clearly show that there is a second chance to make first impressions. However, with regard to warmth, overall ratings decreased between points of measurement. Interestingly, this was only due to a decrease of perceived warmth for the robot-like agent. For the human-like agent ratings remained constant. We can thus conclude to prefer human-like agents as they provide the potential for stable impressions of warmth. Moreover, with regard to competence, employing virtual agents with gestures helps to increase participants' ratings—independent of the agents' appearance. So we can advise to endow virtual agents with gestural behavior to improve their perceived competence. Notably, it should be kept in mind that the gestures employed here were generated from an individualized gesture model which has been rated positively in a previous evaluation study [5]. So there is still the challenge to choose an adequate gesture profile to generate gestures from. Nevertheless, our findings are in line with results from a previous study in which a different virtual agent was employed.

We started with the question whether there is a second chance for first impressions and presented a study that was novel in the way that it combined agent appearance and agent behavior of IVAs as between-subject factors with a repeated measures design to investigate warmth and competence evaluations by human raters. Overall, we aimed to control the experimental setting as much as possible. Nevertheless, one should be aware of the fact that our results are obtained from employing specific IVAs in a specific setting and domain of application (task-related monologue, visual descriptions). The current research points to important issues that need to be studied in future research. These should, on the one hand, isolate further variables like the agents' voice or other nonverbal behaviors, and on the other hand, broaden the scope by investigating different tasks or settings like 'real' interactive dialogues or long-term development. Nevertheless, we are confident that our findings provide an important step for further IVA research in at least two ways. First, we showed that timing actually matters. Results clearly indicate participants' impressions of IVAs change over time providing interesting and detailed information about how we can improve our virtual agents. Second, our methodology of examining specific variables (i.e. agent appearance and agent behavior) allowed for, instead of taking the agent with all its characteristics as a whole, has been shown to be an adequate means to gain detailed insight into the way IVAs are judged by humans.

Acknowledgements This research is partially supported by the Deutsche Forschungsgemeinschaft (DFG) in the Collaborative Research Center 673 "Alignment in Communication" and the Center of Excellence 277 in "Cognitive Interaction Technology" (CITEC). The authors further acknowledge Françoise Therese Rasoloarinala for collecting the data analyzed in this paper.

References

1. Bailenson, J.N., Swinth, K., Hoyt, C., Persky, S., Dimov, A., Blascovich, J.: The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence* 14(4), 379–393 (2005)
2. Bente, G., Krämer, N.C., und Jan Peter de Ruiter, A.P.: Computer animated movement and person perception: Methodological advances in nonverbal behavior research. *Journal of Nonverbal Behavior* 25(3), 151–166 (2001)
3. Bergmann, K., Kopp, S.: GNetlc—Using Bayesian decision networks for iconic gesture generation. In: Ruttkay, Z., Kipp, M., Nijholt, A., Vilhjalmsón, H. (eds.) *Proceedings of the 9th Intern. Conf. on Intelligent Virtual Agents*, pp. 76–89. Springer, Berlin (2009)
4. Bergmann, K., Kopp, S.: Increasing expressiveness for virtual agents—Autonomous generation of speech and gesture in spatial description tasks. In: Decker, K., Sichman, J., Sierra, C., Castelfranchi, C. (eds.) *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems*. pp. 361–368. Budapest, Hungary (2009)
5. Bergmann, K., Kopp, S., Eyssel, F.: Individualized gesturing outperforms average gesturing—evaluating gesture production in virtual humans. In: Allbeck, J., Badler, N., Bickmore, T., Pelachaud, C., Safonova, A. (eds.) *Proceedings of the 10th Conference on Intelligent Virtual Agents*. pp. 104–117. Springer, Berlin/Heidelberg (2010)

6. Bickmore, T., Cassell, J.: Social dialogue with embodied conversational agents. In: van Kuppevelt, J., Dybkjaer, L., Bernsen, N. (eds.) *Advances in Natural, Multimodal Dialogue Systems*. Kluwer Academic Publishers, New York (2005)
7. Brooks, R.: Humanoid robots. *Communications of the ACM* 45(3), 33–38 (2002)
8. Buisine, S., Martin, J.C.: The effects of speech-gesture cooperation in animated agents' behavior in multimedia presentations. *Interacting with Computers* 19, 484–493 (2007)
9. Cassell, J., Thórisson, K.: The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence* 13, 519–538 (1999)
10. Cuddy, A.J., Glick, P., Beninger, A.: The dynamics of warmth and competence judgments, and their outcomes in organizations. *Research in Organizational Behavior* 31, 73–98 (2011)
11. Dautenhahn, K.: Robots as social actors: Aurora and the case of autism. In: *Proceedings of the Third International Cognitive Technology Conference* (1999)
12. Fiske, S.T., Cuddy, A.J., Glick, P.: Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Science* 11(2), 77–83 (2006)
13. Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A., Sasse, M.A.: The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In: *Proceedings of ACM Conference on Human Factors in Computing Systems*. pp. 529–536 (2003)
14. Goetz, J., Kiesler, S., Powers, A.: Matching robot appearance and behavior to tasks to improve human-robot cooperation. In: *Proceedings of the 12th IEEE Workshop on Robot and Human Interactive Communication* (2003)
15. Heylen, D., van Es, I., Nijholt, A., van Dijk, B.: Experimenting with the gaze of a conversational agent. In: *Proceedings International CLASS Workshop on Natural, Intelligent and Effective Interaction in Multimodal Dialogue Systems*. pp. 93–100 (2002)
16. Hinds, P.J., Roberts, T.L., Jones, H.: Whose job is it anyway? a study of human-robot interaction in a collaborative task. *Human Computer Interaction* 19(1), 151–181 (2004)
17. Hoffmann, A., Krämer, N., Lam-Chi, A., Kopp, S.: Media equation revisited. Do users show polite reactions towards an embodied agent? In: Ruttkey, Z., Kipp, M., Nijholt, A., Vilhjálmsson, H. (eds.) *Proceedings of the 9th Intern. Conf. on Intelligent Virtual Agents*. pp. 159–165. Springer, Berlin (2009)
18. Kelley, H.: The warm-cold variable in first impressions of persons. *Journal of Personality* 18, 431–439 (1950)
19. Komatsu, T., Kurosawa, R., Yamada, S.: Difference between users' expectations and perceptions about a robotic agent (adaptation gap) affect their behaviors. In: *Proceedings of the HRI 2011 workshop Expectations in intuitive human-robot interaction* (2011)
20. Krämer, N., Tietz, B., Bente, G.: Effects of embodied interface agents and their gestural activity. In: Rist, T., Aylett, R., Ballin, D., Rickel, J. (eds.) *Proceedings of the 4th International Workshop on Intelligent Virtual Agents*. pp. 292–300. Springer, Berlin/Heidelberg (2003)
21. Krämer, N.C.: *Soziale Wirkungen virtueller Helfer–Gestaltung und Evaluation von Mensch-Computer Interaktion*. Verlag W. Kohlhammer, Stuttgart (2008)
22. Mori, M.: *The Buddha in the robot*. Charles E. Tuttle Co., Tokyo (1982)
23. Niewiadomski, R., Demeure, V., Pelachaud, C.: Warmth, competence, believability and virtual agent. In: Allbeck, J., Badler, N., Bickmore, T., Pelachaud, C., Safonova, A. (eds.) *Proceedings of the 10th Conference on Intelligent Virtual Agents*. pp. 272–285. Springer, Berlin/Heidelberg (2010)
24. Nishio, S., Ishiguro, H.: Attitude change induced by different appearances of interaction agents. *International Journal of Machine Consciousness* 3(1), 115–126 (2011)
25. Peplau, L., Taylor, S., Sears, D.: *Social Psychology*. Prentice Hall (2005)
26. Rehm, M., André, E.: Informing the design of agents by corpus analysis. In: Nishida, T., Nakano, Y. (eds.) *Conversational Informatics*. John Wiley & Sons, Chichester, UK (2007)

27. Schröder, M., Trouvain, J.: The German text-to-speech synthesis system MARY: A tool for research, development and teaching. *International Journal of Speech Technology* 6, 365–377 (2003)
28. Skowronski, J., Amady, N. (eds.): *First Impressions*. The Guilford Press (2008)
29. Sträßling, N., Fleischer, I., Polzer, C., Leutner, D., Krämer, N.C.: Teaching learning strategies with a pedagogical agent: The effects of a virtual tutor and its appearance on learning and motivation. *Journal of Media Psychology: Theories, Methods, and Applications* 22(2), 73–83 (2010)
30. Woods, S., Dautenhahn, K., Schulz, J.: The design space of robots: Investigating children's views. In: *Proceedings of the IEEE Workshop on robot and human interactive communication*. pp. 47–52 (2004)