



Article

A Two-Study Approach to Explore the Effect of User Characteristics on Users' Perception and Evaluation of a Virtual Assistant's Appearance

Carolin Straßmann * and Nicole C. Krämer

Social Psychology: Media and Communication, University Duisburg-Essen, 47057 Duisburg, Germany;
nicole.kraemer@uni-due.de

* Correspondence: carolin.strassmann@uni-due.de

Received: 30 July 2018; Accepted: 22 September 2018; Published: 2 October 2018



Abstract: This research investigates the effect of different user characteristics on the perception and evaluation of an agent's appearance variables. Therefore, two different experiments have been conducted. In a $3 \times 3 \times 5$ within-subjects design (Study 1; $N = 59$), three different target groups (students, elderly, and cognitively impaired people) evaluated 30 different agent appearances that varied in species (human, animal, and robot) and realism (high detail, low detail, stylized shades, stylized proportion, and stylized shade with stylized proportion). Study 2 ($N = 792$) focused on the effect of moderating variables regarding the same appearance variables and aims to supplement findings of Study 1 based on a 3×5 between-subjects design. Results showed effects of species and realism on person perception, users' liking, and using intention. In a direct comparison, a higher degree of realism was perceived as more positive, while those effects were not replicated in Study 2. Further on, a majority evaluated nonhumanoid agents more positively. Since no interaction effects of species and realism have been found, the effects of stylization seem to equally influence the perception for all kind of species. Moreover, the importance of the target group's preference was demonstrated, since differences with regard to the appearance evaluation were found.

Keywords: human–agent interaction; visual appearance; age differences; user characteristics; realism; species

1. Introduction

There is a continuous development of virtual agents and their possible applications. Virtual agents are computer interfaces mostly represented through a virtual embodied character and are humanlike in the way they communicate by using verbal and nonverbal cues [1]. Based on these abilities, those agents are promoted as future personal everyday-life assistants, which might help scheduling and remembering appointments, inform about the weather, or assist with shopping [2]. Although the latest commercial agents (such as Siri or Alexa) are solely speech-based and represented by a voice only, an embodied character was found to enhance the human–agent interaction in a positive way [3]. A meta-analysis revealed the impact of an embodied representation [3]: a comparison of 46 papers showed that the representation of a humanoid face had, in most cases, a more positive outcome compared to the absence of a represented face [3]. Hence, the human–agent interaction seems to benefit from an embodied character.

Furthermore, multiple studies demonstrate that the appearance of those embodied characters affect different social variables (e.g., motivation [4] or the agent's general evaluation [5]). Thus, a virtual agent's appearance has a high impact on the human–agent interaction and is of great importance.

Accordingly, it is important to adapt the appearance of the agent to the user's needs and preferences in order to enhance beneficial effects of human–agent interaction. With regard to the

manifold application fields, the target groups and potential users are diverse and will differ in multiple variables. The application as a virtual assistant, for example, seems to be especially beneficial for people in need of support, such as elderly or cognitively impaired people [6]. Nowadays, those people are dependent on other people's help, but for some tasks, virtual agents can assist their daily life to enable a more self-determined life. This special target group largely differs from digital natives, such as students, who are usually employed in user studies (e.g., regarding technical skills, prior experiences with technologies, or cognitive abilities). Therefore, it is important to analyze the specific needs and preferences of older and/or cognitively impaired users not only with regard to functions, but also with regard to appearance variables. Besides specific target groups, also users with special characteristics concerning attitudes, personality traits, or prior experiences will interact with a virtual assistant. These characteristics also need to be taken into account when the virtual agent's appearance will be analyzed. This might eventually help to enhance the acceptance of these new technologies and to enrich the social interaction.

The best opportunity to tailor the appearance to the user's needs seems to be the customization by the users themselves. However, this process is expensive, time-consuming, and mostly requires technical skills, which special target groups might be afraid of and not able to do. Similarly, results showed that the free customization and own design process of a virtual agent's appearance does not lead to higher user satisfaction compared to a choice out of different options [7]. Therefore, it can be more helpful to know the potential user group and its specific characteristics (e.g., personality traits such as tendency to anthropomorphize) and tailor the appearance or the choice of appearances to these needs. In order to do so, the impact that user characteristics such as age, gender, cognitive impairments, or personality traits have on the perception and evaluation of different appearance variables need to be analyzed. Until now, only little has been known about the preferences of these special target groups such as elderly or cognitively impaired people [8,9] or the influence of user characteristics such as personality traits on the perception of appearance variables. Prior research [8,9] used qualitative approaches to analyze the preferences of elderly people in regard to a virtual assistant's appearance. Although this gives interesting insights, quantitative research is needed to specify these findings more. Additionally, no study investigated the preferences of cognitively impaired people, who are a highly relevant user group [6]. The current approach aims to fill this research gap and explores the effect of user characteristics on the perception and evaluation of different appearance variables in two studies.

Based on previous research, the species of an agent and the degree of realism were identified as key variables to affect human-agent interaction. Accordingly, prior research demonstrates that especially for a specific target group such as elderly, those appearance variables are the ones of most interest [9] and thus need to be investigated in more depth.

Although most embedded characters are humanoid [8], a humanoid character does not necessarily have to be the best choice, since animals, for instance, were found to enhance buying intention [10] or learning outcomes [11] more than humanoid characters. Therefore, it is highly important to take the species of the virtual agent into account more. However, most studies did not systematically investigate the effect of species, since most of the used stimulus materials did also differ regarding other variables such as realism. Realism, in itself, has been shown to be relevant not only in terms of uncanny valley effects [12], but it was also demonstrated that a cartoon stylization can have positive effects on the social interaction [3] and the agent's overall evaluation such as appeal or perceived friendliness [13]. However, realism has not been investigated systematically in combination with the species of the agent. Thus, there is a lack of controlled, systematic research on the effect of species and realism and their potential interactions.

Therefore, the present studies aim to examine the perception and evaluation of these variables in more depth and additionally explore the effects of user characteristics.

2. Theoretical Background

2.1. Effect of a Virtual Agent's Appearance in Human–Agent Interaction

Humans show social reactions towards computers and virtual agents [14,15]. During the interaction with technological entities, humans apply social categories, such as gender stereotypes and ethnical identification, to technological devices (Caucasian faces presented on a computer screen trigger different expectations than Korean faces) and show politeness and reciprocity as would be expected in human–human interaction [16]. Those findings exemplarily highlight that humans apply the same rules they will use in human–human interaction also towards technologies. In this sense, Sobieraj and Krämer [17] showed that users apply the same principles for attractiveness evaluation to a virtual agent's face as they would do for humans. When it comes to an embodied character, the visual appearance contains much more factors (e.g., perceived realism, species of the character, or properties such as height and color) than a human appearance does, since its embodied appearance is fictional and computer-generated. As a result, although humanoid attractiveness rules are also applied to humanlike virtual agents [17], the potential influencing appearance variables are broader and more manifold than those of a human. Therefore, the effects of other variables than only those typical for human appearances have to be investigated.

In order to examine those potential variables in a systematic way, prior research [9] categorized these multiple appearance factors into four main categories (species, realism, 2D vs. 3D, and feature specifications) with several subcategories (see Figure 1). Based on this categorization, the effects of specific appearance variables can be investigated more systematically as appearances can be classified into these categories and only variables of interest receive focused attention.

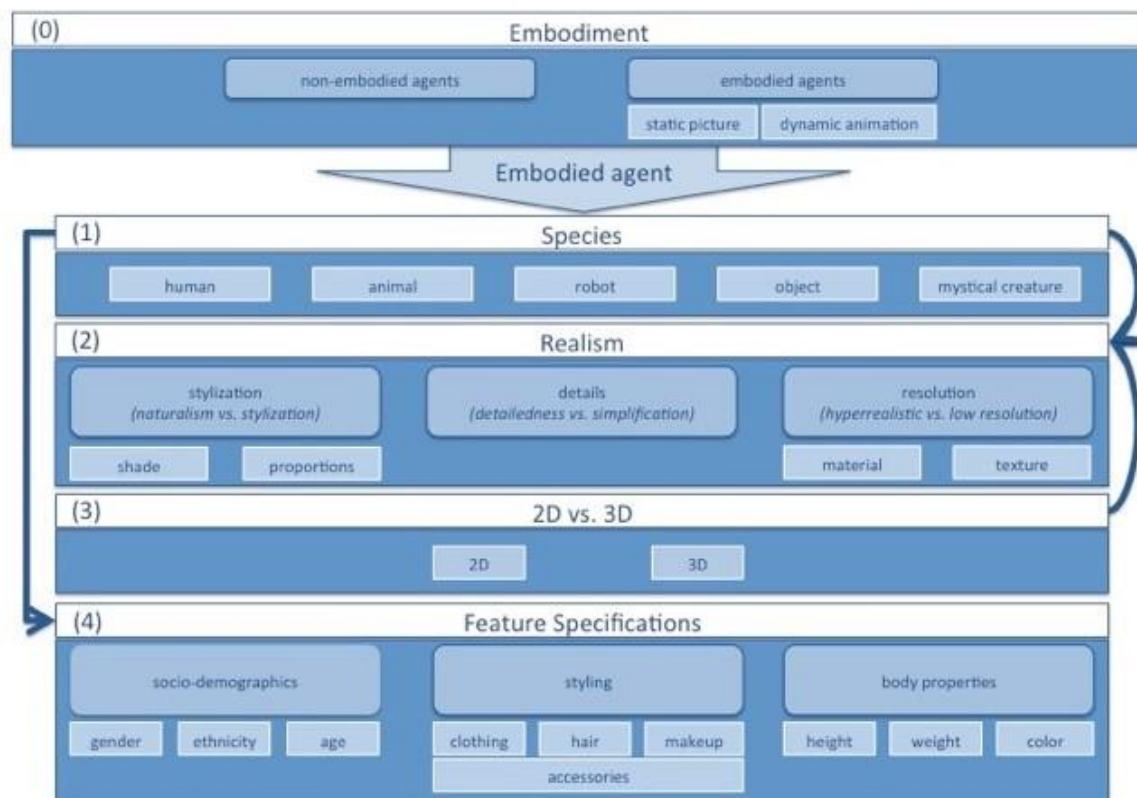


Figure 1. Categorization of different appearance variables. (Reproduced from Straßmann and Krämer [9]).

However, various potential appearance variables and combinations of variables still exist. Since not all of those variables can be considered in an experimental study, the categories of highest

interest need to be investigated more deeply. According to prior research, the categories of species and realism were found to be the ones of most interest [9] and were already found to affect human–agent interaction, e.g., as shown in [5,10,11].

Therefore, the current research focuses on the species and the realism of the agent.

2.2. The Effect of the Agent's Species

Although virtual agents are fictional characters and could easily take up different forms of species (e.g., animals, objects, or humans), Chattaraman et al. [8] found that 91% of virtual agents used on e-commerce websites were humans or humanoid, and findings highlight a preference for humanoid faces [18–20]. However, humanoid agents might not necessarily be the best choice for all occasions. Terada et al. [10] demonstrated that buying intention varies with regard to different species of the agent (Buddha, robot, real person, text, dog, virtual human). It has to be noted, however, that to some degree, the authors not only compared different species, but also their embodiment (text only vs. embodied character) and the perceived realism (real vs. virtual human). The results of that study indicate that an animal and a virtual human evoked the highest values of buying intention, while no embodiment, a robot, and a photorealistic human did not differ from each other and the object leads to the lowest buying intention. Although a virtual human achieved a more positive evaluation than a robot or an object, the animal was also evaluated as being equally positive to a cartoon-like human. Thus, humanoid agents are not necessarily the gold standard. In line with this, a nonhumanoid tutor has been found to be more effective in an online learning environment. Sträfling et al. [11] compared the effect of an humanoid agent to a cartoon-stylized rabbit on learning outcomes. Their results showed that learners preferred a cartoon-like rabbit over a humanoid agent to be their tutor. The rabbit was rated as friendlier and learners exposed themselves longer to the material compared to those who received advice from a humanoid agent. Those results demonstrate that a humanoid character is not always the best solution, since objects or animals have been shown to evoke more engagement and a more positive outcome. However, those results are based on specific appearances, which have not been varied systematically, since the used agents not only differed in species, but also in other appearance variables such as realism. Although these findings provide initial hints that the species of the agent is important and that nonhumanoid appearances seem to have positive effects, further systematic research has to be made to scrutinize the effects of those variables on user responses more closely. Therefore, this research aims to answer the following research question in a more systematic manner:

RQ1: Which species will be perceived and evaluated as most positive?

2.3. The Effect of the Agent's Degree of Realism

As mentioned before, the realism of the agents was also found to have an impact in human–agent interaction. Yee and colleagues [3] conducted a meta-analysis and found that a human-like representation with higher realism leads to more positive social interactions compared to representation with lower realism. However, the meta-analysis showed that this effect was only true for subjective measures and not for behavioral ones. It has to be noted that the authors only used division into high and low levels of realism, without any further distinction. As the categorization of appearance variables (pictured in Figure 1; [9]) demonstrated, realism is more multilayered. Subcategories such as stylization, degree of detail, and resolution affect the overall perception of realism, and thus this concept is more complex than the distinction into high and low realism. Since different subcategories can affect the degree of realism [9] and the meta-analysis gives no further information about these categories, no deductions about the exact effects of these variables can be made. However, other authors have investigated the effects of realism in greater detail and focused on specific realism variables. These studies stressed the effect of more specific realism variables, such as stylization [5,13,21]. McDonnell et al. [13] compared 10 different rendering styles (ranging from line drawing to a highly realistic representation) and found that those which were highly abstract or highly realistic were

perceived as equally appealing. A rendering style that was neither abstract nor realistic was less appealing to the users. Furthermore, this study showed that a cartoon stylization evokes the attribution of friendliness. These findings are in line with the results of Sträfling et al. [11], who found that a cartoon-like rabbit was perceived as more friendly. Similarly, Ring et al. [5] examined the effect of a cartoon-looking shading style and character proportions with regard to different tasks of the virtual agent. Results revealed that a realistic shade was evaluated as being more realistic, while a cartoon shade was rated as being more friendly. In line with this, the cartoon proportions evoked more friendliness than human proportions. Moreover, they found interaction effects of appearance and the presented tasks. A realistic character seems to be more suitable for medical tasks, while the cartoon-like agents were preferred in social tasks [5]. In contrast to this, in the context of health applications, a more realistic-looking agent was evaluated more positively with regard to likability, competence, and trustworthiness compared to a stylized agent [22]. These findings might be caused by the specific application, since Ring and colleagues [5] showed that users prefer a realistic character for medical tasks.

In summary, prior research [5,11,13] showed that a cartoon stylization of a virtual agent evokes the perception of friendliness and that this is true for both the shade (or rendering style) of the agent and the proportions of the agent [5]. At the same time little evidence exists to demonstrate that for medical and health-related tasks a more realistic appearance might be preferred [22].

Nevertheless, it is still unclear how realism variables affect the perception of a virtual agent, especially in the context of daily-life support. As the context of the application seems to be crucial [5], effects might be different for assistive technologies in everyday life. Hence, this research investigates the following research question:

RQ2: What degree of realism will be perceived and evaluated as being most positive in the context of daily-life support?

Since most of the prior studies either focused on species or realism or confounded both, no conclusion of possible interaction effects between both can be made. The stylization of humans might evoke effects that are different from the stylization of other species. Since robots or talking animals are per se artificial and fictional, a stylization of humans might make more of a difference regarding the evaluation.

Prior studies mostly used humanoid agents [5,13,22], while Sträfling et al. [11] showed that a cartoon-like rabbit evoked more friendliness. Therefore, the beneficial effect of a cartoon stylization seems to not necessarily be limited to humanoid agents. However, it has to be mentioned that so far, to the authors' knowledge, no research exists where the effect of different realism variables for varying species has been investigated systematically. As Sträfling et al. [11] only compared a cartoon-like rabbit to a more realistic humanoid agent, it is not derivable whether the positive effect on friendliness is based on the cartoon-like stylization or the species of the agent. Therefore, studies investigating the effect of realism combined with the agent's species are needed.

To fill this research gap, the two studies of the current research aim to give the first insights on these potential interaction effects, by answering this research question:

RQ3: Are there any interaction effects of species and realism with regard to the person perception of the agent, users' liking, and usage intention?

3. Study 1

3.1. Outline and Deviation of Hypotheses

Since various application fields result in different target groups and those target groups might differ in their preferences and needs, Study 1 focuses on target group-related differences (namely, the preferences of seniors and cognitively impaired people in comparison to students) in the perception and evaluation of appearance variables. Modern technologies give great opportunities to support

elderly or cognitively impaired people in their daily life and to maintain their autonomous living as long as possible [2]. Virtual assistance systems are, for example, able to remind people, help them to structure their daily life, or support them with making appointments. It has already been demonstrated that people in need of support accept virtual assistants [6]. These findings are mainly focused on elderly people, but might also be valuable for other target groups, such as cognitively impaired people, as they have similar needs of support. Vardoulakis et al. [23] designed a social relational agent for elderly people (based on a Wizard of Oz setting) in order to care for their mental health and investigate different topics that users may like to talk about with a conversational virtual agent. After talking for one week with the agent on a daily basis, participants stated high levels of companionship, support, and satisfaction and felt comfortable having a virtual agent to talk to at their homes. Moreover, it has been shown that even if interaction problems occur during interaction with a speech-based conversational agent, elderly people and people with cognitive impairments state good levels of acceptance [6].

In summary, an embodied virtual agent is assumed to enhance the daily-life assistance of elderly people and people with special needs. However, those findings are based on interactions with agents that have one specific appearance and thus cannot be generalized for all agents. Since appearance was found to affect the perception and evaluation of a virtual agent in multiple ways, it needs to be investigated how these target groups perceive and evaluate different appearance variables. Thus, what remains open is the question of how a virtual agent for these target groups should look. There are some insights with regard to e-commerce. The results of focus groups with 25 elderly people indicate that abstract agents were preferred, because they are less distracting than humans or humanoid agents [8]. Although these results provide initial insights, it is unclear whether they are transferrable to other applications and tasks. Since in the context of e-commerce, the presented products need to be in the user's focus, a humanoid agent might be distracting; however, in the context of personal assistance, a humanoid agent could be perceived as more serious and engaging. A qualitative interview study in the context of daily-life assistance gave first hints that age-related differences with regard to appearance preferences exist [9]. In contrast to the findings of Chattaraman et al. [8], seniors stated in these interviews that they would prefer a more realistic humanoid agent, while students rejected this kind of appearance [9]. It seems as if seniors, who are mostly less experienced with technologies, strive for higher familiarity than students in order to foster their trust in the agent and to remove boundaries of technology usage. However, both studies used qualitative methods and therefore their results are difficult to generalize. The current study aims to investigate these target groups and their preferences in more detail. Based on the findings of Straßmann and Krämer [9], we hypothesize the following:

Hypothesis 1 (H1). *Seniors evaluate humanoid agents more positively compared to nonhumanoid agents (animals and robots).*

Hypothesis 2 (H2). *Students evaluate nonhumanoid agents (animals and robots) more positively than humanoid ones.*

Hypothesis 3 (H3). *Seniors and students evaluate varying degrees of realism differently.*

Since Yaghoubzadeh et al. [6] showed that virtual assistance is also accepted by cognitively impaired people and as this technology might be highly beneficial for these people, it is important to investigate their special needs. To the authors' knowledge, there are no findings reported to date about the preferences of people with cognitive impairments regarding appearance variables. Nonetheless, it has been shown that people with intellectual disabilities preferred more simple visual representations of hyperlinks to browse more easily through the internet [24]. Consequently, it can be assumed that people with cognitive impairments do also prefer a more simple visual appearance (e.g., a cartoon stylization or reduced detail) to avoid distraction. At the same time, the same principles that are assumed for seniors might be applied, in the sense that this target group requires higher familiarity to enhance acceptance and trust. The current study explores the perception of appearance

variables by people with cognitive impairments for the first time and intends to answer the following research question:

RQ4: How do cognitively impaired people perceive and evaluate different species and degrees of realism?

3.2. Method

The first study aims at investigating the effect of the target group on the evaluation of different appearance variables. Therefore, three different target groups (students, seniors, and cognitively impaired people) participated in this study. A mixed design with one between-subjects factor (3 target groups) and two within-factors (3 species \times 5 realism) was used and all participants evaluated 30 different virtual agents, which differed mainly in regard to species and realism (see Figure 2). The order of the presentation was randomized to prevent sequence effects. The questionnaire was programmed with the tool SoSciSurvey [25]. Participants were invited to the lab, since not all target groups have internet access and to avoid high drop-out rates. The local ethics committee approved the study.

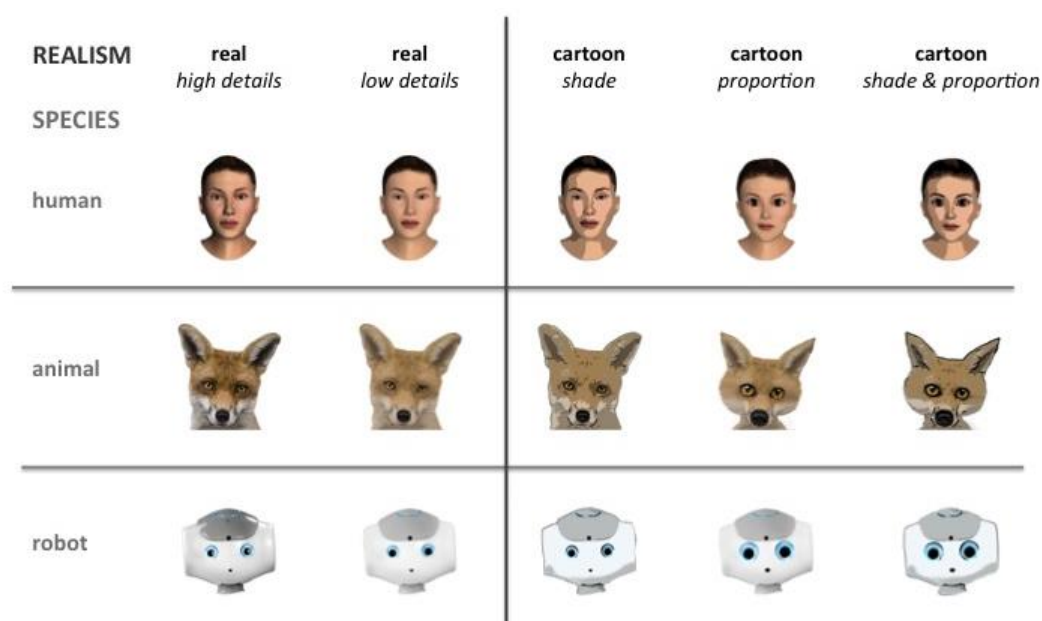


Figure 2. Examples of the used stimulus material.

3.2.1. Sample

Overall, 59 participants from three different target groups completed the questionnaire; 22 students (12 female; age: $M = 21.45$, $SD = 3.99$), 21 seniors (14 female; age: $M = 68.14$, $SD = 8.42$), and 16 people with cognitive impairments (3 female; age: $M = 45.81$, $SD = 19.33$) took part. Half of the participants (31 participants: 11 students, 12 seniors, and 8 cognitively impaired people) knew what a virtual agent is, but only 13 persons (8 students, 4 seniors, and 1 cognitively impaired person) had interacted with one before. The students' data were collected at a large European university, while the cognitively impaired people were recruited in a European Health Care foundation and seniors participated partly in both places. Students were incentivized by course credits, while elderly and cognitively impaired people received 7 Euro as a compensation for their participation. All participants classified as seniors needed to be at least 65 years old. For the cognitively impaired sample, no age restriction was used, but participants were all associated to a psychosocial care facility and were chosen by skilled staff of the Health Care Foundation. No concrete diagnoses were provided for these participants, as on the one hand, this is often very difficult to derive even for skilled persons, and on

the other hand, we wanted to explore the preferences of this generally heterogeneous group. Since a virtual assistant should be useful and effective for people with all kinds of impairments and needs of support, the goal was not to determine the preferences of people with a specific impairment. During the recruitment process, it was checked that those people have a certain degree of cognitive impairments, but that the participants, on the other hand, were able to express their own opinion correctly.

3.2.2. Stimulus Material

As the stimulus material, 30 different static pictures were used. All pictures had the same size and showed the virtual characters' head, which varied in species and realism.

With regard to species, humans, animals, and robots were chosen. Since even within this subcategory, the range of possible appearances is enormous, a pretest with a total of 18 agents (6 for each species) was conducted first. Overall, 24 people (18 female, age: $M = 32.46$, $SD = 12.12$) who were not participants of the two main studies evaluated these agents in a between-subjects design on single items measuring perceived realism, likability of the agent, and willingness to interact. To guarantee generalizability of the results, two agents for each species have been chosen (for example pictures, see Figure 2) based on their likability scores from the pretest. Therefore, to ensure that obtainable differences in the main studies would be caused by the manipulation only, stimuli should not differ in perceived likability, realism, and willingness to interact. In addition, those stimuli with the highest likability scores were chosen. In the end, a woman, a man (both created with Autodesk's character builder: [26]), a fox, a giraffe, the Nao robot, and a more anthropomorphic robot were used. However, the results of the pretest demonstrated that the agents within the species were not evaluated radically differently. Therefore, both versions of the species were collapsed for the calculations.

Realism was manipulated in 5 different degrees based on prior research [5,27]. As mentioned above, multiple subcategories influence the overall perception of realism [9] and therefore are assumed to influence the overall perception of the agent. To investigate these effects in more detail, two subcategories were manipulated: degree of detail and stylization. Thus, the manipulation can be divided into realistic appearances and stylized appearances with a more cartoon-like look. While the realistic appearances are further divisible into high detail and low detail (comparable to the concepts of Gulz and Haake [27]), the stylization was either applied to the proportions, to the shading, or both (see [5]). As realism and its subcategories can be seen as a continuum ranging from high realism to low realism, the created stimuli can be sorted along this continuum: (1) realistic high detail, (2) realistic low detail, (3) stylized shade, (4) stylized proportion, and (5) stylized shade and proportions. With decreasing detail and increasing stylization, the perceived realism of the agent is assumed to decrease as well. The (based on the pretest) chosen appearance pictures all scored medium-high on their resolution. In addition, the appearances had a high level of detail (1), since one can obtain details such as skin or fur properties (for the humanoid agents, wrinkles, rashes, and freckles are displayed, while for the machine-like agents, reflections were obtainable and the fur of the animals was pictured in great detail). To create an appearance with a lower degree of detail (2), those properties were smoothened out. Based on the stimuli with lower detail, the shading (3) was manipulated in a cartoon-like look. Therefore, the outlines were thickened, a soft focus was applied, and the shading itself was manipulated. In order to stylize the agent's proportions (4), facial properties were varied. In dependence of the species, key features such as the eyes and mouth became bigger, while less important parts such as the chin got smaller. Both manipulations (shading and proportions) were applied to create the last degree of realism (5). All manipulations of the degree of realism have been done manually with Photoshop Elements using the same criteria for all species. Figure 2 presents an example of the used stimulus material. In the end, 30 different pictures of agents were created systematically.

3.2.3. Measurements

As a dependent variable, the person's perception was measured repeatedly for each of the 30 stimuli. Five different subscales with an overall 12 items were used: likability, uncanniness, realism, willingness to use, and appeal. Most items have been selected from prior research [13,28], and self-constructed items to measure uncanniness were added. With regard to the within-subjects design and the high number of stimuli, we aimed to keep the measurements short. Therefore, realism, willingness to use, and liking were queried using single items. The scale of likability contained 4 items (not attractive (reversed), unlikable (reversed), reliable, and pleasant) and showed good internal consistency (Cronbach's $\alpha = 0.81$). Furthermore, two items (uncanny and negative) were used to measure uncanniness (Cronbach's $\alpha = 0.84$). Participants rated their agreement to these items on a 5-point Likert scale ranging from 1: "strongly disagree" to 5: "strongly agree". Some additional variables were measured, but since these are not in the focus of the study's research aim, they were not reported in this paper. Since a proportion of the sample had cognitive impairments, the questionnaire was adapted to their special needs in order to guarantee that these participants understand what we intended to measure. The language was adapted in a way that the instructions were simplified, but the items themselves were the same for all three groups to ensure comparability of the results.

3.2.4. Procedure

The experimenter welcomed the participants to the lab and instructed them to fill in the questionnaire at the computer. Before the actual study began, the experimenter informed the participants about the aim and procedure to gain proper informed consent. As done for the measurements, the whole introduction and debriefing material was also adapted to simple language in correspondence with schooled people from the Health Care Foundation. Therefore, proper informed consent was also ensured for people with cognitive impairments. Furthermore, the experimenter assisted the participants (especially those of need in support) whenever it was necessary. However, participants were asked to answer the questionnaire autonomously and informed they should only rely on the experimenter for comprehension problems. After an initial introduction, participants were asked about their prior experiences with virtual agents. Regardless of their answer, a short description of virtual agents was presented to guarantee that all participants had the same definition of a virtual agent in mind. Further on, participants' usage intention was queried. Before the stimulus material was presented, participants were asked to imagine a scenario 5 years in the future, where virtual agents are widely spread. This future scenario was used to ensure that participants were free of thinking about technical restrictions while they were evaluating the presented stimuli. When the main evaluation began, all 30 stimuli were presented in a randomized order. After each stimulus, the dependent variables were assessed. Toward the end, sociodemographic variables were queried. Finally, a debriefing informed the participants about the main research questions.

3.3. Results

To investigate the presented hypotheses and research questions, multiple mixed (Analysis of variance) ANOVAS with a three-component between-subjects factor of target group (students, seniors, and cognitively impaired people) and the two within-subject factors of species (three factors: human, animal, and robot) and realism (high detail, low detail, stylized shade, stylized proportions, and stylized shade with stylized proportions) have been calculated. When the assumption of sphericity was violated, the Greenhouse–Geisser correction is reported. Interaction effects between the within-subject factors and the between-subject factors were further analyzed by calculating the repeated ANOVAs again separately for each target group. In these cases, post-hoc tests using the Bonferroni correction are presented.

3.3.1. Perceived Realism

Appearance effects. First, we examined whether the manipulation was successful and investigated the effect of the manipulation on perceived realism. A significant main effect for species occurred; $F(2, 55) = 37.110, p < 0.001, \eta_p^2 = 0.399$. Post-hoc analyses revealed that all three groups differed significantly from each other. While humans were rated as the most realistic ($M = 3.07, SE = 0.11$), animals ($M = 2.77, SE = 0.12$) were less realistic and robots ($M = 2.00, SE = 0.11$) were least realistic. Further on, a significant main effect for realism emerged; $F(4, 53) = 50.59, p < 0.001, \eta_p^2 = 0.475$. The post-hoc analyses showed that there was no difference between a realistic agent with high and low detail, while the three stylized agents were significantly rated as less realistic than the realistic agent with high and low detail. Furthermore, there was no difference in perceived realism between an agent with a stylized proportion and an agent with a stylized proportion and shading, but a stylized shading was rated as being significantly more realistic compared to both the other stylized agents (see Table 1 for means and standard error).

Table 1. Overall perceived realism ratings for degrees of realism.

Apperances	M	SE
Realistic high detail	3.09 _{a,b,c}	0.10
Realistic low detail	2.98 _{d,e,f}	0.08
Cartoon proportions	2.51 _{a,d,g}	0.11
Cartoon shades	2.28 _{b,e}	0.11
Cartoon proportions and shades	2.22 _{c,f,g}	0.11

Note: Means in columns sharing subscripts are significantly different from each other.

Target-group effects. However, no main effect for target group could be found, indicating that ratings from students, seniors, and cognitively impaired people were, in general, the same regarding the perception of realism.

3.3.2. Likability

Appearance effects. With regard to likability, no main effect of species was found, while a main effect of realism occurred; $F(2.555, 153.093) = 9.861, p < 0.001, \eta_p^2 = 0.150$. Post-hoc analyses revealed that agents with both nonstylized degrees of realism were perceived as significantly more likable than agents with stylized proportions or agents whose shade and proportions were stylized (see Table 2).

Table 2. Likability ratings for agents' degrees of realism of all target groups.

Target Group	Realistic High Detail		Realistic Low Detail		Cartoon Proportions		Cartoon Shades		Cartoon Proportions and Shades	
	M	SE	M	SE	M	SE	M	SE	M	SE
Students	3.28 _{a,b,c}	0.09	3.20 _{d,e,f}	0.1	2.52 _{a,d}	0.1	2.67 _{b,e}	0.09	2.31 _{c,f}	0.1
Seniors	2.96	0.1	2.95 _a	0.11	2.62 _a	0.14	2.77	0.12	2.62	0.13
Cognitively impaired people	2.94	0.14	2.93	0.16	3.28	0.21	3.18	0.21	3.33	0.24
Overall	3.06 _a	0.06	3.02 _b	0.07	2.88	0.08	2.81 _{a,b}	0.08	2.76 _{a,b}	0.09

Note: Means in rows sharing subscripts are significantly different from each other.

Target-group effects. Overall, there was no significant difference in the target groups' likability evaluation. However, significant interaction effects of the species and target group ($F(3.962, 110.938) = 11.151, p < 0.001, \eta_p^2 = 0.285$) as well as of realism and target group ($F(5.110, 143.093) = 16.881, p < 0.001, \eta_p^2 = 0.376$) were found. Post-hoc analyses revealed the following patterns: Students perceived no differences between the three species in likability, while seniors and cognitively impaired people did. Seniors rated the robot ($M = 2.16, SE = 0.12$) as significantly less likable than humans ($M = 3.25, SE = 0.16$) and animals ($M = 2.94, SE = 0.18$). In contrast, cognitively impaired people evaluated robots ($M = 3.53, SE = 0.22$) as being more likable than animals ($M = 2.77,$

SE = 0.16), while they perceived no differences concerning likability between robots and humans ($M = 3.09$, $SE = 0.24$) or humans and animals. With regard to realism, students rated agents with both nonstylized degrees of realism as being significantly more likable than the three stylized versions, while seniors only saw differences in likability between the realistic version with low detail and the stylized proportions. Cognitively impaired people perceived agents with all degrees of realism as being equally likable. (Consult Table 2 for means and post-hoc analyses.)

3.3.3. Uncanniness

Appearance effects. Further calculations showed significant main effects of species ($F(2, 112) = 3.735$, $p = 0.027$, $\eta_p^2 = 0.063$) and realism ($F(2.746, 153.767) = 252.843$, $p < 0.001$, $\eta_p^2 = 0.820$) on perceived uncanniness. Post-hoc analysis demonstrated that animals ($M = 2.12$, $SE = 0.09$) were evaluated as being significantly less uncanny compared to robots ($M = 2.38$, $SE = 0.09$), while humans ($M = 2.26$, $SE = 0.08$) did not differ significantly from animals or robots. Further on, agents with a stylized shade (also true for the combination with a stylized proportion) were perceived as being less uncanny compared to those with a realistic shading (high detail, low detail, and stylized proportions) (see Table 3).

Table 3. Uncanniness ratings for agents' degrees of realism of all target groups.

Target Group	Realistic High Detail		Realistic Low Detail		Cartoon Proportions		Cartoon Shades		Cartoon Proportion and Shades	
	M	SE	M	SE	M	SE	M	SE	M	SE
Students	2.24 _{a,b,c}	0.12	2.45 _{d,e,f}	0.15	3.06 _{a,d,g,h}	0.15	1.46 _{b,e,g}	0.06	1.60 _{c,f,h}	0.07
Seniors	2.67 _{a,b}	0.13	2.58 _{c,d}	0.14	2.79 _{e,f}	0.13	1.41 _{a,c,e}	0.07	1.47 _{b,d,f}	0.08
Cognitively impaired people	3.18 _{a,b}	0.22	3.12 _{c,d}	0.23	2.86	0.24	1.47 _{a,c,e}	0.12	1.42 _{b,d,f}	0.12
Overall	2.70 _{a,b}	0.09	2.72 _{c,d}	0.1	1.45 _{a,c,e}	0.05	2.90 _{e,f}	0.1	1.50 _{b,d,f}	0.05

Note: Means in rows sharing subscripts are significantly different from each other.

Target-group effects. Again, no main effect of target group emerged as being significant. However, significant interaction effects of the species and target group ($F(3.959, 110.853) = 7.792$, $p < 0.001$, $\eta_p^2 = 0.218$) as well as of realism and target group ($F(5.492, 153.767) = 11.292$, $p < 0.001$, $\eta_p^2 = 0.287$) were found. Students again perceived no differences in uncanniness between the species. The same pattern that was found for likability with regard to robots occurred for seniors and cognitively impaired people: seniors perceived robots ($M = 2.64$, $SE = 0.14$) as being more uncanny than humans ($M = 1.90$, $SE = 0.13$) and animals ($M = 2.01$, $SE = 0.14$), while cognitively impaired people evaluated robots ($M = 2.13$, $SE = 0.23$) as being least uncanny (humans: $M = 2.39$, $SE = 0.20$; animals: $M = 2.71$, $SE = 0.16$). Moreover, cognitively impaired people evaluated appearances without any stylization as being more uncanny than students and seniors did (Table 3).

3.3.4. Liking of the Agent

Appearance effects. Further on, we found significant differences between the species with regard to the participants' liking of the agent ($F(2, 112) = 5.232$, $p = 0.007$, $\eta_p^2 = 0.085$). Post-hoc analyses revealed that humans ($M = 2.68$, $SE = 0.11$) were more liked than robots ($M = 2.25$, $SE = 0.11$), while animals ($M = 2.50$, $SE = 0.13$) did not differ from robots or humans. The degree of realism also affected the participants' evaluation of liking ($F(2.957, 165.615) = 19.350$, $p < 0.001$, $\eta_p^2 = 0.257$). The lower the degree of realism, the less people liked the appearance.

Agents with a realistic rendering (high and low detail) were significantly more liked than appearances with a cartoon stylization (all three conditions), while an appearance with cartoon-stylized proportions evoked significantly higher ratings of liking than the combination of a stylized shading and stylized proportions (see Table 4).

Table 4. Liking of the agent for degrees of realism from all target groups.

Target Group	Realistic High Detail		Realistic Low Detail		Cartoon Proportions		Cartoon Shades		Cartoon Proportion and Shades	
	M	SE	M	SE	M	SE	M	SE	M	SE
Students	2.92 _{a,b,c}	0.14	2.74 _{d,e}	0.17	2.12 _{a,d}	0.14	2.37 _{b,f}	0.12	1.97 _{c,e,f}	0.12
Seniors	2.22	0.16	2.12	0.16	1.9	0.17	2.07	0.15	1.94	0.16
Cognitively impaired people	3.04	0.2	3.08	0.2	2.99	0.23	2.83	0.19	2.77	0.25
Overall	2.73 _{a,b,c}	0.1	2.65 _{d,e,f}	0.1	2.43 _{a,d,g}	0.08	2.34 _{b,e}	0.1	2.23 _{c,f,g}	0.1

Note: Means in rows sharing subscripts are significantly different from each other.

Target-group effects. Target groups did not differ in their evaluation of liking. However, again, interaction effects were obtained for target group and species ($F(3.809, 106.662) = 4.980, p < 0.001, \eta_p^2 = 0.151$) as well as for realism ($F(5.815, 165.615) = 4.964, p < 0.001, \eta_p^2 = 0.151$). Again, students' liking of the species was not significantly different, and seniors liked robots ($M = 1.52, SE = 0.12$) significantly less than humans ($M = 2.37, SE = 0.26$) and animals ($M = 2.26, SE = 0.21$). In contrast to prior findings, cognitively impaired people liked animals ($M = 3.38, SE = 0.17$) more than robots ($M = 2.64, SE = 0.25$), while there was no significant difference between robots and humans ($M = 2.81, SE = 0.24$) or humans and animals. Furthermore, for seniors and cognitively impaired people, the degree of realism had no effect on whether they liked the agent. Students liked more realistic agents more than ones with stylized proportions and proportions and shade (see Table 4).

3.3.5. Willingness to Use

Appearance effects. Moreover, the participants' willingness to use the agent was examined. The mixed ANOVA showed significant main effects for species ($F(2, 112) = 3.798, p = 0.025, \eta_p^2 = 0.064$) and realism ($F(2.916, 163.274) = 16.062, p < 0.001, \eta_p^2 = 0.223$). Post-hoc tests showed that participants were more likely to use a humanoid agent ($M = 2.62, SE = 0.11$) compared to a machinelike one ($M = 2.27, SE = 0.10$). Similar to participants' liking behavior, the usage intention decreased with the degree of realism. Post-hoc tests revealed that appearances with a realistic stylization evoked the highest usage intention values, while participants stated that they were least likely to use an agent with stylized shading combined with stylized proportions (see Table 5).

Table 5. Participant's usage intention of agents for degrees of realism from all target groups.

Target Group	Realistic High Detail		Realistic Low Detail		Cartoon Proportions		Cartoon Shades		Cartoon Proportion and Shades	
	M	SE	M	SE	M	SE	M	SE	M	SE
Students	2.69 _{a,b,c}	0.15	2.51 _{d,e,f}	0.16	2.01 _{a,d}	0.13	2.06 _{b,e,g}	0.12	1.80 _{c,f,g}	0.11
Seniors	2.06	0.16	2.1	0.16	1.91	0.17	2.02	0.17	1.93	0.16
Cognitively impaired people	3.19	0.21	3.19	0.24	2.99	0.22	3.08	0.21	2.97	0.24
Overall	2.65 _{a,b,c}	0.1	2.60 _{d,e,f}	0.11	2.39 _{a,d,g}	0.1	2.30 _{b,e}	0.1	2.23 _{c,f,g}	0.1

Note: Means in rows sharing subscripts are significantly different from each other.

Target-group effects. Further on, a significant main effect of target group was found ($F(2, 56) = 735.755, p < 0.001, \eta_p^2 = 0.929$). Cognitively impaired people ($M = 3.08, SE = 0.17$) stated higher values of usage intention compared to seniors ($M = 2.01, SE = 0.15$) and students ($M = 2.21, SE = 0.15$). Additionally, interaction effects of the manipulations and the target groups were found (species: $F(3.73, 104.54) = 4.047, p = 0.005, \eta_p^2 = 0.126$; realism: $F(5.83, 163.27) = 4.913, p < 0.001, \eta_p^2 = 0.149$). The same patterns that were shown with regard to liking behavior occurred for usage intention: for students, the species had no effect on their usage intention, while seniors liked to use robots ($M = 1.52, SE = 0.13$) significantly less than humans and animals (humans: $M = 2.23, SE = 0.27$; animals: $M = 2.27, SE = 0.21$) and cognitively impaired people showed higher usage intention for animals ($M = 3.38, SE = 0.18$) compared to robots ($M = 2.88, SE = 0.23$). With regard to realism, students

stated a higher usage intention for agents with a realistic stylization, while seniors and cognitively impaired people showed no differences with regard to realism in their usage intention (see Table 5).

3.4. Interim Conclusions

In line with prior research, the presented findings give hints that species and realism affect the agent's evaluation and user's usage intention of the agent. While there was no main effect of species on the perceived likability, robots were evaluated as being more uncanny compared to animals. However, humanoid agents were more liked and evoked higher usage intention compared to machinelike agents. Although appearances with stylized shading were perceived as being less uncanny than those with realistic shading, a higher degree of realism was rated as being more likable than a stylized appearance. Overall, the liking of the agent and user's usage intention decrease with the degree of realism.

Results of Study 1 suggest that overall, there are differences between the target groups. For students, the species of the agent seems to be of lower importance, since they evaluated the agent's person perception of the different species equally. Overall, seniors rejected robots and rated them more negatively than humans and animals. Cognitively impaired people rated robots as being more likable and less uncanny than other species, but liked animals more and stated higher usage intention for animals. In contrast, the degree of realism has mostly no effect on person perception for seniors and cognitively impaired people, while students evaluated a more realistic appearance as being more positive. Thus, results indicate that students seem to rely more on realism, while the species is more important for seniors and cognitively impaired people.

4. Study 2

4.1. Outline and Deviation of Hypotheses

Results of Study 1 give the first insights into the effect of species and realism on an agent's perception and evaluation, especially in regard to target group-related differences. Nevertheless, this study has some limitations that make it difficult to derive generalizable implications. Due to the different target groups and the high effort of recruiting this group of people, the sample size of Study 1 was rather small and a within-subjects design was used. Therefore, these characteristics might have influenced the results and produced more differences between stimuli.

Study 2 therefore employs a between-subjects design in order to complement the findings of Study 1. Instead of focusing on differences between specific user groups, the current study mainly investigates the effects of species (RQ1), realism (RQ2), and their potential interaction effect (RQ3) in a more controlled manner with a larger sample. In addition, to further contribute to the question of user characteristics, different moderating variables are taken into account to investigate the influence of personality traits. This is based on research which shows that personality traits have an impact on the evaluation of an artificial entity [29].

One user characteristic of interest is the tendency to anthropomorphize, which can be defined as "the tendency to apply human characteristics (i.e., emotions, motivations, and goals) to nonhuman animals, objects, and natural entities" [30] (p. 214). It has already been shown that anthropomorphism tendency is positively related to the perception of uncanniness [31]. However, these insights were found in the context of the perception of readers of fiction, and thus it has to be investigated whether they are also valid for the perception of artificial entities. However, since both are fictional, the same principles might well be applicable, and it can be therefore assumed that the tendency to anthropomorphize affects the perception of different appearance variables. People who have a higher tendency to anthropomorphize might feel more comfortable to interact with a fictional character, since they are more able to attribute humanoid characteristics to nonhumanoid characters. Therefore, we assume the following:

Hypothesis 4 (H4). *A higher tendency to anthropomorphize leads to more liking of nonhumanoid agents.*

In the context of human–robot interaction, negative attitudes and anxieties towards these technologies were found to affect the interaction [32]. We assume that these effects are also transferable to virtual agents. People with negative attitudes or anxieties toward agents might evaluate varying appearance variables differently. For instance, people who are afraid of social harm caused by virtual agents might be more likely to wish for a clear distinction between the virtual and real world and therefore prefer a higher degree of stylization and more artificial species.

Hypothesis 5 (H5). *Negative attitudes and anxieties towards virtual agents moderate the effect of species and realism on liking and usage intention.*

4.2. Method

For Study 2, an online study with a 3 (species: human, animal, and robot) \times 5 (degree of realism: high detail, low detail, cartoon proportions, cartoon shade, and cartoon proportions and shade) between-subjects design has been conducted. The questionnaire was accepted for the recruitment of participants by the SoSciSurvey panel [33]. This panel has about 65,000 active members. While the age is almost balanced, most of those people in the panel have a high level of education. The local ethics committee approved the study.

4.2.1. Sample

Overall, 792 people filled in the questionnaire. Gender was not equally balanced, since 304 (39%) men and 471 (59%) women participated. In addition, 17 people (2%) did not want to state their gender. Age ranged from 15 to 80 years, with an average age of 39 years ($M = 38.63$, $SD = 14.77$). When participants had interacted with an agent before, most of them had talked to an agent without an embodiment (78%).

4.2.2. Stimulus Material

Appearance was manipulated in the same way as in Study 1. In this way, species were varied between humans, animals, and robots (with two variations of each species), which were manipulated in five different degrees of realism (high detail, low detail, stylized shade, stylized proportions, and stylized shade and proportions). Thus, the same 30 pictures (see. Figure 2) were used, but this time, in a between-subjects design. Every participant saw and evaluated only one picture.

4.2.3. Measurements

As dependent variables, participants evaluated the stimulus material regarding person perception, liking, perceived usefulness, usage intention, and trust. Person perception was measured with five different subscales: anthropomorphism, likability, appeal, trustworthiness, and competence. These were five-point semantic differentials and most of the items originated from the Goodspeed Questionnaire [28] and the measures of McDonnell et al. [13]. Anthropomorphism (e.g., unreal–real or machine-like–humanoid) and likability were measured with six pairs of adjectives, while appeal (e.g., appealing–not appealing or attractive–unattractive), trustworthiness (e.g., trustworthy–not trustworthy or reliable–not reliable), and competence (e.g., competent–incompetent, intelligent–not intelligent) contained four item pairs. Liking was a self-constructed scale with five items (e.g., if I had a personal virtual agent, I would wish that the agent would look exactly like this) to measure how much participants liked the presented appearance. To measure usage intention, perceived usefulness, and trust, a scale from Heerink et al. [34] was transferred to the application of virtual agents. Participants rated the items of those scales on a five-point Likert scale ranging from 1: “strongly disagree” to 5: “strongly agree”. All dependent variables showed good or excellent internal consistency (see Table 6).

Table 6. Dependent variables' reliability values.

Concept	Number of Items	Cronbach's Alpha
Anthropomorphism	6	0.836
Likability	6	0.950
Appeal	4	0.916
Trustworthiness	4	0.810
Competence	4	0.875
Liking	5	0.887
Usage intention	3	0.914
Perceived usefulness	3	0.926
Trust	2	0.920

To measure the users' tendency to anthropomorphize things, a scale with 10 items (e.g., I sometimes wonder if my computer deliberately runs more slowly after I have shouted at it) was used [30]. From the original scale with 20 items, only those 10 items focusing on the present behavior and feelings were used. The internal consistency was good (Cronbach's alpha = 0.82). The users' attitude toward agents (three items, e.g., I think it's a good idea to use a virtual agent) and anxiety towards agents (four items, e.g., if I should use a virtual agent, I would be afraid to make mistakes with it) originates from Heerink et al. [34]. These scales showed an acceptable reliability (Cronbach's alpha = 0.84 and 0.78, respectively). Additionally, the scale of negative attitude towards robots [35] with its three subscales was transferred to the context of virtual agents. Although the Cronbach's alpha values were not sufficient (ranging from 0.59 to 0.69), this scale was used, since it was very relevant for the presented research questions and is a well-established scale in the realm of human–robot interaction. Again, participants used a five-point Likert scale to state their agreement for all moderating variables. Furthermore, participants stated their age, gender, educational background, and prior experiences with virtual agents.

4.2.4. Procedure

In the beginning, participants were welcomed and thanked for their willingness to participate. After this initial introduction, their tendency to anthropomorphize was queried. Thereafter, general questions about virtual agents such as prior experiences or the participant's attitude towards this technology were assessed. Before the picture of the agent was presented, participants were asked to imagine the same future scenario as the one used in Study 1. Each participant saw one of the 30 different agents and rated it with regard to the dependent variables. After the participants' sociodemographics were measured, participants were debriefed. At the end, participants had the chance to take part in a lottery.

4.3. Results

4.3.1. Person Perception

To investigate the effect of species and realism on person perception, a (multivariate analysis of variance) MANOVA with two between-subject factors (species and realism) on anthropomorphism, likability, appeal, trustworthiness, and competence has been calculated. Main effects of species on all variables except for competence were found (see Table 7). All three species differed with regard to anthropomorphism; while animals scored highest, humans reached moderate values and robots were perceived as being least anthropomorphic (Table 8). With regard to likability, appeal, and trustworthiness, animals differed significantly from humans and robots: animals were perceived as more likable, more appealing, and more trustworthy than humans and robots (see Table 8).

Table 7. Results of MANOVA with regard to person perception.

Dependent Variable	Source	df	F	p	η^2
Anthropomorphism	Species	2	88.593	<0.001	0.280
	Realism	4	7.459	<0.001	0.061
	Species X Realism	8	4.611	<0.001	0.075
Error		456			
Likability	Species	2	18.389	<0.001	0.075
	Realism	4	1.745	0.139	0.015
	Species X Realism	8	1.267	0.259	0.022
Error		456			
Attractiveness	Species	2	21.383	<0.001	0.086
	Realism	4	3.493	0.008	0.030
	Species X Realism	8	1.349	0.217	0.023
Error		456			
Trustworthiness	Species	2	9.506	<0.001	0.040
	Realism	4	1.267	0.282	0.011
	Species X Realism	8	1.008	0.429	0.017
Error		456			
Competence	Species	2	1.978	0.139	0.009
	Realism	4	1.947	0.102	0.017
	Species X Realism	8	1.651	0.108	0.028
Error		456			

Table 8. Means of person perception for species.

Species	Anthropomorphism		Likability		Attractiveness		Trustworthiness		Competence	
	M	SD	M	SD	M	SD	M	SD	M	SD
Human	2.21 _a	0.71	2.54 _a	0.86	2.22 _a	0.84	2.40 _a	0.76	2.91	0.81
Animal	2.93 _a	0.95	3.18 _b	1.00	2.90 _{a,b}	0.91	2.81 _{a,b}	0.092	2.95	0.82
Robot	1.87 _a	0.62	2.69 _{a,b}	1.06	2.35 _b	0.97	2.50 _b	0.87	3.09	0.81
Overall	2.32	0.89	2.80	1.02	2.49	1.02	2.57	0.87	2.99	0.82

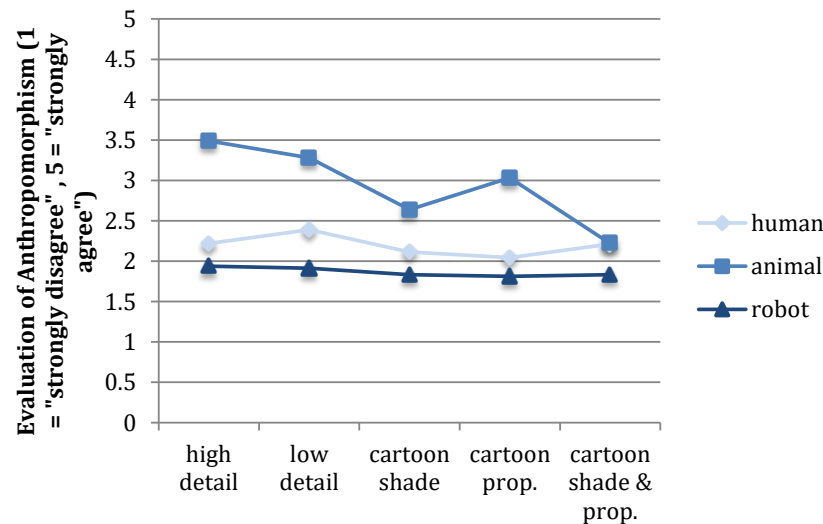
Note: Means in columns sharing subscripts are significantly different from each other.

In addition, a significant main effect of realism occurred for anthropomorphism ($F(4, 456) = 7.459$, $p < 0.001$, $\eta_p^2 = 0.061$) and appeal ($F(4, 456) = 3.493$, $p = 0.008$, $\eta_p^2 = 0.030$). Post-hoc analyses with Bonferroni correction revealed that the stylization of the shade leads to lower anthropomorphism, since a stylized shade (also in combination with a stylized proportion) differed significantly from both versions without any stylization (high and low detail) (see Table 9). In addition, the realistic style with lower detail was perceived as being more appealing than the complete cartoon stylization. Interaction effects of both independent variables were only found for anthropomorphism (Table 7). Animals were seen as more anthropomorphic than humans and robots for most of the realism degrees, but when the cartoon stylization was applied to both proportion and shade, no difference between the species occurred (Figure 3).

Table 9. Means of person perception for realism.

Degree of Realism	Anthropomorphism		Likability		Attractiveness		Trustworthiness		Competence	
	M	SE	M	SE	M	SE	M	SE	M	SE
High detail	2.55 _{a,b}	0.08	2.86	0.10	2.60	0.10	2.58	0.09	2.92	0.08
Low detail	2.53 _{c,d}	0.07	2.99	0.10	2.74 _a	0.09	2.72	0.08	3.18	0.08
Cartoon shade	2.20 _{a,c}	0.08	2.71	0.11	2.37	0.11	2.51	0.09	2.90	0.09
Cartoon proportion	2.30	0.08	2.68	0.10	2.41	0.10	2.47	0.09	2.95	0.09
cartoon shade and proportion	2.09 _{b,d}	0.08	2.74	0.10	2.30 _a	0.10	2.55	0.09	2.97	0.08

Note: Means in columns sharing subscripts are significantly different from each other.

**Figure 3.** Interaction effect of species and realism on perceived anthropomorphism.

4.3.2. Liking of the Agent

Moreover, a two-factorial ANOVA with species and realism as independent variables and liking as the dependent variable showed a main effect only for species ($F(2, 734) = 7.578, p = 0.001, \eta_p^2 = 0.020$), indicating that animals ($M = 1.96, SD = 0.96$) and robots ($M = 1.94, SD = 0.99$) were more liked than humans ($M = 1.68, SD = 0.74$). However, no main effect for realism occurred, nor was an interaction effect found.

4.3.3. Usage Intention, Perceived Usefulness, and Trust

A second MANOVA has been calculated with usage intention, perceived usefulness, and trust towards virtual agents as dependent variables and species and realism as independent variables. We found significant main effects of species (usage intention: $F(2, 744) = 1.769, p = 0.008, \eta_p^2 = 0.013$ and perceived usefulness: $F(2, 744) = 6.899, p = 0.001, \eta_p^2 = 0.018$), while no effect of realism nor an interaction effect was revealed. Usage intention was higher for robots than for humans and animals, while robots also were perceived as being more useful than animals (Table 10).

Table 10. Participant's usage intention, perceived usefulness, and trust ratings for all agent species.

Species	Usage Intention		Perceived Usefulness		Trust	
	M	SD	M	SD	M	SD
Human	1.92 _a	1.04	2.31	1.12	1.91	0.93
Animal	1.91 _b	1.03	2.18 _a	1.05	1.87	0.94
Robot	2.16 _{a,b}	1.11	2.53 _a	1.13	1.97	0.94
Overall	2.00	1.06	2.34	1.10	1.92	0.94

Note: Means in columns sharing subscripts are significantly different from each other.

4.3.4. Moderating Variables

To investigate the effect of user characteristics such as age, gender, anthropomorphic tendency, attitudes toward agents, anxieties toward agents, and the negative attitude toward agents on usage intention and liking of the agent, two hierarchical regression analyses were calculated separately. Predictors were entered in the following order: age and gender (step 1); the tendency to anthropomorphize (step 2); and attitude toward agents, anxieties toward agents, and all three subscales of negative attitude toward agents (step 3). As presented in Table 11, the anthropomorphism tendency and participants' attitude toward agents significantly explains the variance of usage intention and liking of the agent. The greater the tendency to anthropomorphize or the more positive the participant's attitude toward agents, the higher the participant's usage intention and liking of the agent. Moreover, the anxiety toward agents and negative attitudes toward the social influence of agents significantly contribute to the explanation of the variance of usage intention (Table 11). The usage intention increases with anxiety towards agents and decreases with increasing negative attitudes towards the social influence of agents.

Table 11. Hierarchical multiple regression analyses.

Steps of the Regression Analyses	Usage Intention					Liking of the Agent				
	<i>b</i> (<i>SE_b</i>)	β	<i>t</i>	<i>p</i>	ΔR^2	<i>b</i> (<i>SE_b</i>)	β	<i>t</i>	<i>p</i>	ΔR^2
Step 1					0.000					0.005
Gender	0.01 (0.11)	0.00	0.08	0.934		0.09 (0.10)	0.05	0.90	0.368	
Age	0.00 (0.00)	0.02	0.40	0.690		0.00 (0.00)	0.06	1.10	0.271	
Step 2					0.037					0.038
Anthropomorphism Tendency	0.38 (0.10)	0.19	3.72	<0.001		0.35 (0.10)	0.20	3.71	<0.001	
Step 3					0.280					0.070
Anxiety toward agents	0.19 (0.08)	0.16	2.53	.012		0.15 (0.08)	0.14	1.97	0.050	
Attitude toward agents	0.43 (0.07)	0.40	6.61	<.001		0.17 (0.07)	0.17	2.45	0.015	
Negative attitude toward situations of interaction with agents	−0.15 (0.09)	−0.11	−1.66	.097		−0.11 (0.09)	−0.09	−1.17	0.244	
Negative attitude toward social influence of agents	−0.19 (0.08)	− 0.14	−2.55	.011		−0.14 (0.08)	−0.12	−1.87	0.062	
Negative attitude toward emotions in interaction with agents	−0.11 (0.08)	−0.09	−1.45	0.148		−0.05 (0.08)	−0.04	−0.56	0.579	
Total <i>R</i> ²					0.317					0.112
Step1		<i>F</i> (2, 358) = 0.08, <i>p</i> = 0.924					<i>F</i> (2, 353) = 0.84, <i>p</i> = 0.431			
Step 2		<i>F</i> (3, 357) = 4.68, <i>p</i> = 0.003					<i>F</i> (3, 352) = 5.18, <i>p</i> = 0.002			
Step 3		<i>F</i> (8, 352) = 20.46, <i>p</i> < 0.001					<i>F</i> (8, 347) = 5.48, <i>p</i> < 0.001			

Note: Values in bold indicate significant relationships.

To investigate a possible moderating influence of anthropomorphism tendency, attitudes toward agents, and anxieties toward agents on the effects of both independent variables (species and realism) on participants' liking and usage intention, multiple moderation analyses using the Hayes process were calculated separately. The effects of the potential moderating variables are presented in Figures 4 and 5. Against our predictions, all 12 moderation analyses revealed no moderation effects of anthropomorphism tendency, attitudes toward agents, and anxieties toward agents in the effect of species or realism on liking and usage intention.

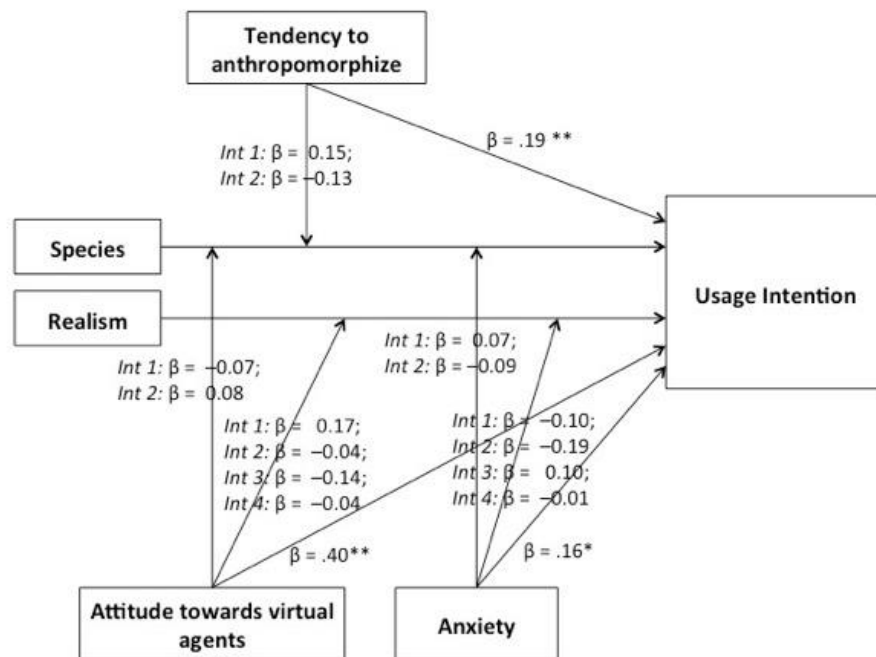


Figure 4. Conceptual model of the moderating variables' effects on usage intention. Standardized regression coefficients are presented (* indicates a p value < 0.005 and ** < 0.001). For the moderation analyses, independent variables were dummy coded (species: Int 1 = human vs. animal, Int 2 = human vs. robot; realism: Int 1 = high detail vs. low detail, Int 2 = high detail vs. cartoon proportions, Int 3 = high detail vs. cartoon shade, Int 4 = high detail vs. cartoon proportions with cartoon shade) and data were standardized before running the analyses.

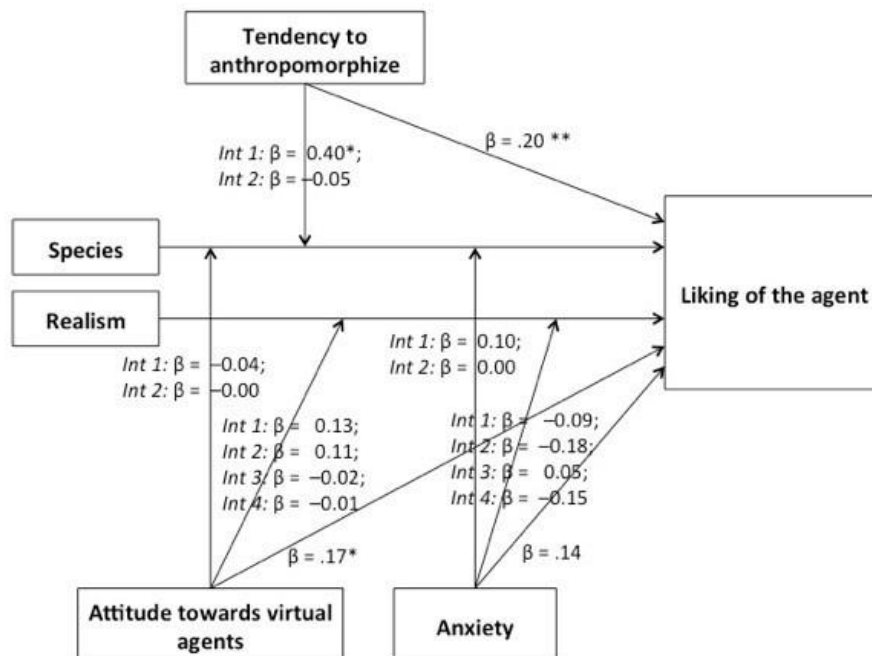


Figure 5. Conceptual model of the moderating variables' effects on liking of the agent. Standardized regression coefficients are presented (* indicates a p value < 0.005 and ** < 0.001). For the moderation analyses, independent variables were dummy coded (species: Int 1 = human vs. animal, Int 2 = human vs. robot; realism: Int 1 = high detail vs. low detail, Int 2 = high detail vs. cartoon proportions, Int 3 = high detail vs. cartoon shade, Int 4 = high detail vs. cartoon proportions with cartoon shade) and data were standardized before running the analyses.

4.4. Interim Conclusion

All three species vary in their perceived degree of anthropomorphism, in the sense that animals were perceived as most anthropomorphic, humans reached moderate anthropomorphism ratings, and robots were evaluated as least anthropomorphic. Additionally, animals were seen as more likable, more appealing, and more trustworthy than humans and robots. While animals and robots were more liked than humanoid agents, participants' usage intention was higher for robots than for animals and humans.

The stylization of the shade leads to lower anthropomorphism ratings, while a realistic agent with low detail was evaluated as being more appealing than an agent with stylized shades and proportions. The degree of realism had no effect on likability, liking of the agent, or usage intention.

Regression analyses revealed that participants' tendency to anthropomorphize as well as their general attitude towards agents predicts the liking of the agent and participants' usage intention. However, no moderating effects were found for user characteristics.

5. General Discussion

In the context of human-agent interaction, the appearance of the agents was found to have main effects. Prior research demonstrates that the species (e.g., human or animal) and realism (e.g., cartoon shade) of the agent affect the agent's evaluation [5,10,11,13] and is therefore important for the outcome of the human-agent interaction. However, until now, both factors had not been examined in a controlled and systematic way. Therefore, we varied three different species (human, animal, and robot) and five degrees of realism (realistic-style high detail, realistic-style low detail, cartoon-stylized shade, cartoon-stylized proportions, and complete cartoon stylization) within two different studies. We investigated how different species (RQ1) and degrees of realism (RQ2) were evaluated and whether there is an interaction effect of both variables (RQ3). Since many applications have special target groups (e.g., people in need of support), we further investigated the preferences of different user groups (H1, H2, H3, and RQ4) and the influence of further user characteristics such as personality traits (H4 and H5). When the influences of the target group and user characteristics are examined, the appearance can be tailored to the needs of the users. This will enhance the interaction and acceptance of virtual agents. In the following, the results of both studies are summarized and discussed based on prior research.

5.1. Effects of Species (RQ1)

With regard to species and research question 1 (RQ1), both studies present contradicting findings. While Study 1 showed that users preferred humans to robots and showed higher usage intention for humans, this was not the case in Study 2. People seem to prefer humanoid agents in a direct comparison (as in the within-subjects design of Study 1), but this is not true when only one agent is presented (as in the between-subjects design of Study 2). The between-subjects design showed that animals were perceived as more likable, appealing, and trustworthy than humans and robots. Moreover, participants liked animals and robots more than humans, and those species evoked higher usage intention. As the results of Study 2 represent a wider sample (no specific target groups and higher sample size) and these results are more generalizable than those of Study 1, nonhumanoid agents seem to be evaluated as being more positive than humanoid ones. However, most of the agents employed in current systems are humanoid. Based on our findings, the gold standard of providing a humanoid agent needs to be reconsidered. While for specific target groups such as senior people, humanoid agents might be more appropriate, the results of Study 2 emphasize that at the same time, a majority were found to like nonhumanoid agents more than humanoid ones.

5.2. Effects of Realism (RQ2)

The investigation of research question 2 (RQ2) also led to contradictory findings. Study 1 demonstrates that participants liked appearances with a higher degree of cartoon stylization less than those with a realistic style. The same pattern was obtained for the usage intention of the participants. Furthermore, agents with a realistic stylization were evaluated as being more likable. However, these findings were not replicated in our second study, since the degree of realism had no effect on likability, liking of the agent, or participants' usage intention. Previous studies mostly showed that lower degrees of realism evoke a more positive perception of the agent [10,13,36]. However, there is prior research that supports findings of Study 1, since Van Wissen et al. [22] showed that participants preferred a realistic-looking agent over a cartoon-stylized agent to be their virtual nurse. The most reasonable explanation for these results seems to be the type of application, since Ring et al. [5] and also van Wissen et al. [22] showed that realistic agents are more appropriate for medical tasks. In line with this, Robertson et al. [37] found that people even stated that they get angry about the use of cartoon-stylized agents in medical applications, since it somehow stultifies a very serious task. In our studies, the task itself was not explicitly defined, but participants were asked to imagine the agents as their own personal assistant in everyday life. Thus, this application field might underlie the same principles as a medical task, since both applications aim to provide support and help. Therefore, an appearance with a higher degree of realism seems to be more appropriate. The discrepancy between both studies might be caused by the different designs, since Study 1 used a within-factor design, while participants in Study 2 only evaluated one agent. Realism seems to be more important when users are able to compare different appearances. Since the different realism degrees were manipulated in five small steps (relying on different subcategories), the differences might be much clearer in a direct comparison. While one can easily obtain the agent's species (and the perception associated therewith), the realism of the agent is more complex and multilayered and its degrees are more subtle. Therefore, differences might be stronger in a within-subject design.

5.3. Interaction Effects of Species and Realism (RQ3)

Furthermore, the present studies examined the interaction effects of species and realism for the first time in a systematic manner (RQ3). However, nearly no interaction effects were found (with the exception of perceived anthropomorphism). It can be concluded that the stylization of an agent has similar effects on all kinds of species (such as humans, robots, and animals). In line with this, prior research showed that animals with atypical features such as enlarged eyes were rated as less familiar [38], which is also true for humanoid agents with stylized proportions. While in Study 1, the effects of realism degrees were bigger (higher effect sizes), in Study 2, effects of the agent's species were stronger. Thus, as described above, the degrees of realism seem to be more important in a direct comparison (e.g., when users are able to choose from alternatives), while the species was more decisive in a between-subjects design. When people only see one agent, the species of the agent is more decisive for the impression-building process.

5.4. Effects of the Target Group (H1, H2, H3, H4, H5, and RQ4)

This was based on initial hints that indicated that there are age-related differences with regard to appearance preferences [9]. Based on these findings, we assumed that seniors evaluate humanoid agents more positively than nonhumanoid ones (H1), while students might rate nonhumanoid agents more positively (H2). This assumption was found partly to be true with regard to the evaluation of robots. While students liked robots the most and showed highest values of usage intention for robots, seniors clearly rejected machinelike agents. These findings are in line with the results of prior qualitative research [9]. In interviews, seniors stated that they would prefer a humanoid agent, since the interaction with it is more familiar as they talk to humans all the time, while animals or other species cannot answer appropriately [9]. However, these results cannot be generalized to all people in need of

support, since cognitively impaired people evaluated robots more positively and stated higher usage intention and liking for animals; this group seems to prefer nonhumanoid agents more. While for seniors, a high familiarity seems to be of importance and only a humanoid appearance is appropriate as their assistant, cognitively impaired people are less restricted to humans. The cognitively impaired people are probably used to receiving assistance in their daily life, while seniors mostly were more afraid to eventually be in need of support. Therefore, for seniors, it might be more important to have something familiar and serious as a humanoid agent, while cognitively impaired people are a little more open-minded in the sense that they are willing to accept assistance from different species. Differences between those target groups were also found in relation to realism. Thus, hypothesis H3 was confirmed. However, in contrast to the findings of Straßmann and Krämer [9], seniors evaluated a cartoon stylization more positively, while students rated realistic agents more positively. The overall patterns of Study 1 indicate that for people in need of support (seniors and cognitively impaired people), the species is more important for their evaluation, liking, and intention to use. In contrast, for students, the degree of realism was more important than it was for seniors and cognitively impaired people. Results emphasize that there is no universal appearance factor that is appropriate for various user groups.

5.5. Effects of User Characteristics (H4 and H5)

Overall, no moderating effects of personality or attitudes were found. Therefore, hypothesis H4 and H5 need to be rejected. Nevertheless, anthropomorphism tendency and participants' attitude toward virtual agents predict the liking of the agent and usage intention. While those variables were found to not moderate the effects of appearance variables, they influence whether people are willing to interact with a virtual agent and whether they like it. People with a high tendency to anthropomorphize and a more positive attitude toward virtual agents liked the agent more. Furthermore, the attitude and anxiety towards virtual agents affect the willingness to use it. Thus, to enhance the interaction with virtual agents, people's attitudes need to be improved. Although we found no moderating effects with regard to appearance, future research should investigate whether appearance variables can have a positive effect on people's attitudes or decrease anxieties toward agents in general.

6. Limitations and Future Work

For both studies, several limitations have to be taken into account. One major limitation is caused by the presentation of pictures only. Since participants evaluated the agent based on static pictures, only limited generalizations to user behavior in real human-agent interactions can be made. Although the used method was also beneficial, since participants could not get distracted by the agent's behavior or interaction characteristics and therefore evaluated appearance factors only, studies with agents that interact with the users are needed. Thus, future studies should investigate whether the presented findings are transferable to user behavior in real interactions and whether there is still an influence of appearance on perceptions.

Even though the stimulus material was prepared in a more systematic way than in previous studies and was based on a pretest, the stimuli are still characterized by other appearance variables beside species and realism (e.g., styling of humanoid agents, chosen animals, or colors). To minimize the effect of those variables, two versions for each species have been chosen and evaluations have been collapsed, which leads to a higher generalizability of the findings. In addition, as mentioned above, the degrees of realism that were manipulated were relatively similar. Although the very systematic manipulation of the appearance factors is a big strength of the present research, with regard to our findings, it might also be a limitation. As only differences in realism in a within-subjects design have been found, it has to be mentioned that differences are very subtle and might only be obtainable in a direct comparison. Despite the fact that the participants were instructed to imagine a future scenario in order to reduce mental restrictions, e.g., with regard to technical implementations, participants might still have been limited in their imagination.

Study 1 aimed to investigate the preferences of different target groups and especially focused on people in need of support. Therefore, elderly and cognitively impaired people participated. It has to be noticed that some participants of the cognitive impairment condition could also be seen as seniors, since their age was over 65. However, the distinction has been made based on participants' cognitive skills, which were evaluated by schooled employees of an established Health Care Foundation. Since people with different sorts of impairments—which had not been specified, except for the fact that those people had low cognitive skills—participated, findings are limited and could not be generalized to specific impairment groups (e.g., autistic people). Since Study 1 indicates that people with cognitive impairments, in general, have specific preferences, future studies could focus on specific impairments to obtain deeper insights into the preferences and perceptions of this target group. In addition, the sample size of the cognitively impaired subsample was relatively low, since the recruitment of these participants is not easy. Therefore, the findings of this study can rather be seen as initial hints. However, since virtual assistance is highly beneficial for those users, our research is a valuable start in the investigation of the needs of this target group.

The within-subjects design of Study 1 also has to be reflected critically, since the evaluation of all 30 pictures took almost one hour for some of the participants. Especially for people with cognitive impairments, the participation most probably was very exhausting, although the language of the questionnaire was adapted to the participants' skills to lower the cognitive effort. In order to avoid fatigue, the experimenter helped the participants and gave them a break whenever it was needed. However, it might be beneficial to replicate the effects of the target group in a between-subjects design.

In the present research, two studies have been conceptualized to supplement each other, but the results of both studies are contradictory rather than supporting. Possible explanations have been addressed in the discussion part, but nevertheless, further studies are needed to clarify these contradicting findings. Although such contradictions might decrease comprehensiveness of the implications, with regard to open science and the need for replications of scientific findings, contradicting results might occur more often. Overall, these results highlight that multiple (even more than two) studies are always needed to achieve reliable insights.

7. Conclusions

This research stresses the importance of different appearance variables, such as species and realism. Both factors were found to influence the users' person perception and preferences with regard to usage intention and perceived usefulness. Effect sizes revealed that in a direct comparison of different appearance alternatives, the realism seems to be more important. However, for the impression perception of one agent without any comparisons, the agent's species is more decisive. Results emphasize that a majority evaluated nonhumanoid agents more positively and therefore might be highly beneficial in human–agent interaction. Although animals or robots evoked a more positive perception in general when a wider sample was gathered, they might not be suitable for all kinds of user groups, as seniors were found to clearly reject robots as their personal assistant. These findings help to understand the preferences of specific target groups more deeply and therefore will be beneficial when designing appropriate appearances of daily-life assistants.

Author Contributions: C.S. and N.K. conceived and designed the experiments; C.S. performed the experiments; C.S. analyzed the data; C.S. wrote the original paper; N.K. reviewed and edited the paper; N.K. supervised.

Funding: This work was part of the project KOMPASS (Socially cooperative, virtual assistants as companions for people in need of cognitive support) (Grant No. 16SV7272) that is funded by the German Federal Ministry of Education and Research.

Acknowledgments: In this section you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g. materials used for experiments).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Cassell, J.; Prevost, S.; Sullivan, J.; Churchill, E.; Van Mulken, S.; Klesen, M.; Baldes, S. The automated design of believable dialogues for animated presentation teams. In *Embodied Conversational Agents*; MIT Press: Cambridge, MA, USA, 2000; pp. 220–255.
2. Kopp, S.; Brandt, M.; Buschmeier, H.; Cyra, K.; Krämer, N.; Kummert, F.; Opfermann, C.; Pitsch, K.; Schillingmann, L.; Straßmann, C.; et al. Conversational Assistants for Elderly Users—The Importance of Socially Cooperative Dialogue. In Proceedings of the AAMAS 2018 Workshop on Intelligent Conversation Agents in Home and Geriatric Care Applications, Stockholm, Sweden, 15 July 2018.
3. Yee, N.; Bailenson, J.N.; Rickertsen, K. A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces. In Proceedings of the Conference on Computer-Human Interaction (CHI), San Jose, CA, USA, 28 April–3 May 2007.
4. Baylor, A.L. Promoting motivation with virtual agents and avatars: Role of visual presence and appearance. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2009**, *364*, 3559–3565. [[CrossRef](#)] [[PubMed](#)]
5. Ring, L.; Utami, D.; Bickmore, T. The Right Agent for the Job? The Effects of Agent Visual Appearance on Task Domain. In Proceedings of the 14th International Conference on Intelligent Virtual Agents, Boston, MA, USA, 27–29 August 2014; pp. 374–384.
6. Yaghoubzadeh, R.; Kramer, M.; Pitsch, K.; Kopp, S. Virtual Agents as Daily Assistants for Elderly or Cognitively Impaired People. In Proceedings of the 13th International Conference on Intelligent Virtual Agents, Edinburgh, UK, 29–31 August 2013; pp. 79–91.
7. Diehl, C.; Schiffhauer, B.; Eyssel, F.; Achenbach, J.; Klett, S.; Botsch, M.; Kopp, S. Get One or Create One: The Impact of Graded Involvement in a Selection Procedure for a Virtual Agent on Satisfaction and Suitability Ratings. In Proceedings of the International Conference on Intelligent Virtual Agents, Stockholm, Sweden, 27–30 August 2017; pp. 109–118.
8. Chattaraman, V.; Kwon, W.-S.; Gilbert, J.E.; Shim, S.I. Virtual agents in e-commerce: Representational characteristics for seniors. *J. Res. Interact. Mark.* **2011**, *5*, 276–297. [[CrossRef](#)]
9. Straßmann, C.; Krämer, N.C. A Categorization of Virtual Agent Appearances and a Qualitative Study on Age-Related User Preferences. In Proceedings of the International Conference on Intelligent Virtual Agents, Stockholm, Sweden, 27–30 August 2017; pp. 413–422.
10. Terada, K.; Jing, L.; Seiji, Y. Effects of Agent Appearance on Customer Buying Motivations on Online Shopping Sites. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, Seoul, Korea, 18–23 April 2015; pp. 929–934.
11. Sträfling, 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. *J. Media Psychol.* **2010**, *22*, 73–83. [[CrossRef](#)]
12. Mori, M.; MacDorman, K.F.; Kageki, N. The uncanny valley. *IEEE Robot. Autom. Mag.* **2012**, *19*, 98–100. [[CrossRef](#)]
13. McDonnell, R.; Breidt, M.; Bülthoff, H.H. Render me real? *ACM Trans. Graph.* **2012**, *31*. [[CrossRef](#)]
14. Hoffmann, L.; Krämer, N.C.; Lam-chi, A.; Kopp, S. Media Equation Revisited: Do Users Show Polite Reactions towards an Embodied Agent? In Proceedings of the International Conference on Intelligent Virtual Agents, Amsterdam, The Netherlands, 14–16 September 2009; pp. 159–165.
15. Reeves, B.; Nass, C. *How People Treat Computers, Television, and New Media Like Real People and Places*; CSLI Publications and Cambridge University Press: Cambridge, UK, 1996; ISBN 1575860538.
16. Nass, C.; Moon, Y. Machines and Mindlessness: Social Responses to Computers. *J. Soc. Issues* **2000**, *56*, 81–103. [[CrossRef](#)]
17. Sobieraj, S.; Krämer, N.C. What Is Beautiful in Cyberspace? Communication with attractive avatars. In Proceedings of the International Conference on Social Computing and Social Media, Crete, Greece, 22–27 June 2014; pp. 125–136.
18. Forlizzi, J.; Zimmerman, J.; Mancuso, V.; Kwak, S. How interface agents affect interaction between humans and computers. In Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces, Helsinki, Finland, 22–25 August 2007.
19. Parise, S.; Kiesler, S.; Sproull, L.; Waters, K. Cooperating with life-like interface agents. *Comput. Hum. Behav.* **1999**, *15*, 123–142. [[CrossRef](#)]

20. Koda, T.; Maes, P. Agents with faces: The effect of personification. In Proceedings of the 5th IEEE International Workshop on Robot and Human Communication, Tsukuba, Japan, 11–14 November 1996; pp. 189–194.
21. Zell, E.; Aliaga, C.; Jarabo, A.; McDonnell, R. To Stylize or not to Stylize? The Effect of Shape and Material Stylization on the Perception of Computer-Generated Faces. *ACM Trans. Graph.* **2015**. [CrossRef]
22. Van Wissen, A.; Vinkers, C.; van Halteren, A. Developing a Virtual Coach for Chronic Patients: A User Study on the Impact of Similarity, Familiarity and Realism. In *Persuasive Technology*; Meschtscherjakov, A., De Ruyter, B., Fuchsberger, V., Murer, M., Tscheligi, M., Eds.; Springer: Cham, Switzerland, 2016; pp. 263–275.
23. Vardoulakis, L.P.; Ring, L.; Barry, B.; Sidner, C.L.; Bickmore, T.W. Designing Relational Agents as Long Term Social Companions for Older Adults. In Proceedings of the International Conference on Intelligent Virtual Agents, Santa Cruz, CA, USA, 12–14 September 2012; pp. 289–302.
24. Rocha, T.; Bessa, M.; Bastardo, R.; Magalhães, L. Image-type representation: A preliminary study on preferences of users with intellectual disabilities. *Int. J. Hum. Comput. Stud.* **2018**, *110*. [CrossRef]
25. Gulz, A.; Haake, M. Visual Design of Virtual Pedagogical Agents: Naturalism versus Stylization in Static Appearance. In Proceedings of the International Design and Engagability Conference, Salzburg, Austria, 5–7 April 2006; pp. 1–9.
26. SoSci Survey Tool. Available online: <http://www.soscisurvey.de> (accessed on 26 September 2018).
27. Autodesk's Character Generator. Available online: <http://www.charactergenerator.autodesk.com> (accessed on 26 September 2018).
28. Bartneck, C.; Kulić, D.; Croft, E.; Zoghbi, S. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *Int. J. Soc. Robot.* **2009**, *1*, 71–81. [CrossRef]
29. Isbister, K.; Nass, C. Consistency of personality in interactive characters: Verbal cues, non-verbal cues, and user characteristics. *Int. J. Hum. Comput. Stud.* **2000**, *53*, 251–267. [CrossRef]
30. Neave, N.; Jackson, R.; Saxton, T.; Hönekopp, J. The influence of anthropomorphic tendencies on human hoarding behaviours. *Pers. Individ. Differ.* **2015**, *72*, 214–219. [CrossRef]
31. Meullemans, C.G.E.; Abraham, J. The Prediction of Sensitivity towards Uncanniness: An Applied Psychological Study on Novel Readers in the Light of Anthropomorphism and Attitudes towards Thrownness. *IOP Conf. Ser. Mater. Sci. Eng.* **2017**, *180*, 12248. [CrossRef]
32. Nomura, T.; Kanda, T.; Suzuki, T.; Kato, K. Prediction of human behavior in human-robot interaction using psychological scales for anxiety and negative attitudes toward robots. *IEEE Trans. Robot.* **2008**, *24*, 442–451. [CrossRef]
33. Leiner, D.J. Our research's breadth lives on convenience samples—A case study of the online respondent pool "SoSci Panel". *Stud. Commun. Media* **2016**, *5*, 367–396. [CrossRef]
34. Heerink, M.; Kröse, B.; Evers, V.; Wielinga, B. Assessing acceptance of assistive social agent technology by older adults: The almere model. *Int. J. Soc. Robot.* **2010**, *2*, 361–375. [CrossRef]
35. Nomura, T.; Kanda, T.; Suzuki, T. Experimental investigation into influence of negative attitudes toward robots on human-robot interaction. *AI Soc.* **2006**, *20*, 138–150. [CrossRef]
36. Li, J.; Kizilcec, R.; Bailenson, J.; Ju, W. Social robots and virtual agents as lecturers for video instruction. *Comput. Hum. Behav.* **2015**. [CrossRef]
37. Robertson, S.; Solomon, R.; Ridl, M.; Wicklin Gillespie, T.; Chociemski, T.; Master, V.; Mohan, A. The Visual Design and Implementation of an Embodied Conversational Agent in a Shared Decision-Making Context (eCoach). In Proceedings of the International Conference on Learning and Collaboration Technologies, Los Angeles, CA, USA, 2–7 August 2015; Volume 9192, pp. 427–437.
38. Schwind, V.; Leicht, K.; Jäger, S.; Wolf, K.; Henze, N. Is there an Uncanny Valley of Virtual Animals? A Quantitative and Qualitative Investigation. *Int. J. Hum. Comput. Stud.* **2017**, *111C*, 49–61. [CrossRef]

