Int J Soc Robot (2010) 2: 175–186
DOI 10.1007/s12369-010-0056-9

A Cross-cultural Study: Effect of Robot Appearance and Task

Dingjun Li · P.L. Patrick Rau · Ye Li

Accepted: 26 April 2010 / Published online: 13 May 2010

Abstract This study investigates the effects of culture, robot appearance and task on human-robot interaction. We propose a model with culture (Chinese, Korean and German), robot appearance (anthropomorphic, zoomorphic and machinelike) and task (teaching, guide, entertainment and security guard) as factors, and analyze these factors' effects on the robot's likeability, and people's active response to, engagement with, trust in and satisfaction with the robot. We conducted a laboratory experiment with 108 participants to test the model and performed Repeated ANOVA and Kruskal Wallis Test on the data. The results show that cultural differences exist in participants' perception of likeability, engagement, trust and satisfaction; a robot's appearance affects its likeability, while the task affects participants' active response and engagement. We found the participants expected the robot appearance to match its task only in the interview but not in the subjective ratings. Interaction between culture and task indicates that participants from low-context cultures may have significantly decreased engagement when the sociability of a task is lowered. We found strong and positive correlations between interaction performance (active response and engagement) and preference (likeability, trust and satisfaction) in the human-robot interaction.

Keywords Human robot interaction · Social robot · Cultural differences · Robot appearance · Robot task

D. Li · P.L.P. Rau (✉) · Y. Li
Institute of Human Factors and Ergonomics, Department of Industrial Engineering, Tsinghua University, Beijing, P.R. China
e-mail: rpl@mail.thu.edu.cn

D. Li
State Intellectual Property Office of P.R. China, Beijing, P.R. China

1 Introduction

Ranging from professional service robots used for security, medical or logistic purposes to private and personal robots for domestic or entertaining use, service robots are becoming ubiquitous throughout the world. Millions of service robots had been sold by the end of 2008 and the number is expected to more than double over the next four years [1]. However, it should be noted that robot producers generally offer identical robotic products for users in different cultures without varying the design. It is possible that users in these cultures perceive and react to the robots in a way that is contrary to the designers' expectations. As a result, taking full account of the cultural diversity in users' backgrounds becomes highly important for designers of robotic products.

Besides, it is important to match a robot's appearance with its task. Robot producers deliver robots with diverse appearances to perform different tasks, such as machine-like (e.g., SuperDroid's remote presence robot [2]), animal-like (e.g., Sony's AIBO for entertainment [3]), caricatured (e.g., Takara's TERA SECURITY for home security [4]) and humanoid (e.g., Aldebaran Robotics' NAO for research or entertainment [5]). When a robot is delivered and functions in a specific circumstance, the first impression of the robot plays a central role in initiating and maintaining the interaction. As shown in previous studies, people tend to form a mental model and expectations of a robot's qualifications even before the interaction begins [6, 7]. Thus, there is high risk of creating a misleading mental model for the user if the robot's appearance is not properly designed. For example, a doglike robot may be appropriate for entertaining children, but if it were used to deliver lectures to students, they may feel uneasy or consider the lecture material to be unreliable.

Designing a proper robot appearance to match its task is even more complicated when the robot will function in an international environment. People's mental model will not only be influenced by the robot's appearance but will also be mediated by their cultural background.

In our study, we proposed several hypotheses related to the importance of matching a robot's appearance with its task for different cultures, and conducted a laboratory experiment to test these hypotheses. The goal was to identify how cultural backgrounds affect people's reaction to and perception of a robot, and to determine the proper appearance for specific tasks in a cross-cultural context. The results should be useful in the design of service robots that will be operating across cultures.

2 Literature Review

2.1 Cultural Differences

Culture influences every aspect of people's social behavior. Geert Hofstede's cross-cultural study in more than 50 countries provided a framework to explain the different social behaviors with the national culture [8–10]. In his study, five dimensions were identified and labeled "Power Distance", "Uncertainty Avoidance", "Individualism versus Collectivism", "Masculinity versus Femininity" and "Long- versus Short-term Orientation" respectively.

Power Distance (PD) is defined as "the extent to which the less powerful members of organizations and institutions accept and expect that power is distributed unequally". **Uncertainty Avoidance** (UA) indicates "the extent to which a culture programs its members to feel either uncomfortable or comfortable in unstructured situations". Unstructured situations are "novel, unknown, surprising, and different from usual". **Individualism** (IDV) versus collectivism is "the extent to which the ties between individuals are loose: everyone is expected to look after him/herself and his/her immediate family". On the collectivist side, people are integrated into strong in-groups. **Masculinity** (MAS) versus femininity is defined as "the extent to which the dominant values in society are 'masculine'—that is, assertiveness, the acquisition of money and things, and not caring for others and the quality of life". **Long-Term Orientation** (LTO) associates with values of thrift and perseverance whereas Short-Term Orientation associates with values of respect for tradition, fulfilling social obligations and protecting one's 'face' [10].

National cultures could be scored and located on the cultural dimensions. The cultural dimension scores for China, Korea and Germany are illustrated in Fig. 1, where a high score on one dimension indicates the culture is close to one side of the dimension [11]. For example, China has a very high score on the dimension of Long-Term Orientation, which means Chinese society is strongly inclined to time perspective and it has an attitude of persevering.

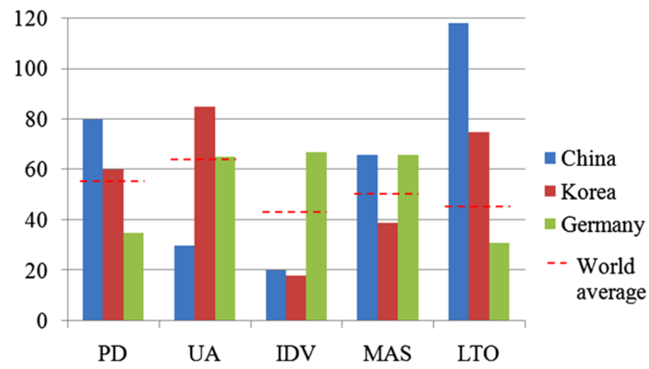

Fig. 1 Hofstede's cultural dimension scores for China, Korea and Germany [11]

In addition, Hall identified the concept of high-context and low-context to categorize communication styles in different cultures. In a high-context culture, people value contextual elements that are not expressed explicitly, whereas in a low-context culture, most information must be in the transmitted message [12]. China and most Asian countries are generally classified as high-context cultures, and Germany is a typical low-context culture [12]. The culture dimensions could be used to explain peoples' different social behaviors and perception across cultures.

2.2 Towards Culturally Adaptive Robots

In 1997, O'Neill-Brown introduced the concept of culturally adaptive agents, which emphasized that the design of intelligent adaptive agents should consider users' cultural background and the cultural context. In her study, O'Neill-Brown argued that culture influences every aspect of people's modes of interaction with agents, such as communication style, personality, knowledge, motivation, cognitive modes, individual and group relationships, mixture of verbal and non-verbal communication and group dynamics [13]. The importance of designing culturally adaptive software has been fully stressed [14]; however, the exploration of culturally adaptive robots is just beginning.

Based on the concept of culturally adaptive agents, we define the *culturally adaptive robot* as a type of social robot that can detect users' cultural cues in communicative behaviors and respond to them appropriately. The aim of a culturally adaptive robot is to promote natural and smooth interaction between user(s) and robot(s). To design such a robot, understanding cultural differences in how people perceive and react to the robot is central.

Recent research on human-robot interaction has confirmed the cultural influence on human-robot interaction in many aspects. Bartneck's early studies indicated the cultural differences in people's negative attitudes towards robots and explained the different attitudes with different levels of exposure to robots [15, 16]. A follow-up study examined cultural differences in people's perceptions of a robot's appearance, which indicated that for US participants, the likeability was greater when the robot was more anthropomorphic, while Japanese participants showed the opposite trend [17]. Although the studies did not provide any in-depth explanation about why the cultural differences existed, they set the stage for research on cultural differences in human-robot interaction.

There is evidence that people prefer and are more responsive to culturally adaptive robots. In Rau, Li et al.'s study, German participants greatly preferred the explicit communication style of a robot (e.g. says "I think this choice is not correct" to express disagreement), while their Chinese counterparts preferred an implicit one (e.g. says "Are you sure?" to show disagreement) [18]. This finding was further confirmed in a subsequent experiment, in which American participants were more influenced by a robot that communicated explicitly, while Chinese participants were influenced more by the implicit one [19]. Moreover, Chinese subjects were more likely to consider a robot as an in-group member than the Americans [19]; research further showed that Chinese participants, compared with American ones, felt more comfortable with an in-group robot assistant [20]. The above studies indicated that to improve users' positive feelings during the interaction, robots must be designed to provide and respond to culturally adaptive social cues corresponding to the users' background.

From the limited studies that involved real human-robot interaction in 3-D space, we found evidence that culturally adaptive robots are correlated with more positive perception and evaluation. However, more in-depth studies are required to form a systematical framework of the design of culturally adaptive robots. Thus, in our study we analyzed how people's cultural backgrounds influence their perception and evaluation of a robot.

2.3 Robot Appearance

Slightly modifying Fong's robot morphology [21], we classify the appearance of robots into three types: anthropomorphic (human-like), zoomorphic (animal-like) and functional (neither human-like nor animal-like, but related to a robot's function) appearance.

Previous studies have revealed that a robot's appearance can affect the user's expectation, perception and evaluation of a robot's behavior and capabilities [22–26]. Masahiro Mori first proposed a theoretical concept named the "Uncanny Valley" to reflect the nature of people's perceptions of and attitudes towards robots. The "Uncanny Valley" illustrated that as a robot's appearance and actions are made more humanlike, people's emotional response will become increasingly positive and empathetic, until a point is reached beyond which a response of revulsion may occur [27]. Gee and Browne explained the upward part of the "Uncanny Valley" with the increasing familiarity towards certain objects [28]. In Walter et al.'s study, participants indicated preference to a robot with some facial features [29] and considered it more intelligent than a robot without any facial features [30], which might be an evidence of the upward part of the "Uncanny Valley". Besides, many more studies have tried to explain the downward part of the "Uncanny Valley", in which a robot is quite humanlike but still different from humans. For example, Sarah Woods's study of children's perception of robots' appearance indicated that children considered the robots with more humanoid features (e.g. a humanoid body shape, eyes, a mouth etc.) to be more aggressive, while they considered the robots with more machine-like features (e.g. wheels, tracks, no facial features etc.) to be friendlier [26]. This study provided some tentative empirical evidence for the downward part of the "Uncanny Valley". Evidence from controlled experimental studies is expected to support or challenge the "Uncanny Valley" theory.

2.4 Compatibility Between Robot Appearance and Task

When considering the task domain, many researchers suggested the consistency between a robot's appearance and its task [23, 25, 31]. In Goetz et al.'s study, they used a 2-D robotic head with three levels of human-likeness (i.e. human, a mixture of human and machine, machine) to analyze people's perception of what the robot could do [23]. In Lohse et al.'s study, the robots were classified as humanoid, zoomorphic and functional robots. Participants were presented short movies with descriptions of robots, and were asked about appropriate tasks for these robots [31]. In both studies, some appropriated applications were identified for each type of robots. Some suggested applications for three types of robots' appearances are listed in Table 1.

The results should be used with caution since participants had shown quite diverse ideas on robots' application and each application was described within a certain context. Besides, both studies used pictures or videos of robots instead of real robots, and no real interactions between the participants and the robots occurred. People's perception of appropriate tasks for a robot might be different if the robot interacts with them in the real world.

3 Research Framework and Hypotheses

We propose a conceptual model of people's cultural background, a robot's appearance and its task to analyze their effect on people's perception of the human-robot interaction. The research framework is shown in Fig. 2. It should be noted that we designed the four tasks to have different levels of sociability (measured by interaction time and frequency):

Table 1 Proposed applications for robots with different appearances

Application		Types of robots' appearances		
Category	Example	Humanoid	Zoomorphic	Functional
Security	Military tasks exploration		x	x
Public assistance	Tour guide translator	x		x
Business	Reception sales	x		
Research	Lab assistant for research	x	x	
Healthcare	Therapy food carrier	x	x	
Education	Language teacher		x	
Entertainment	Toy pet		x	

* The symbol "x" indicates a strong correlation between an application and a type of robots' appearances

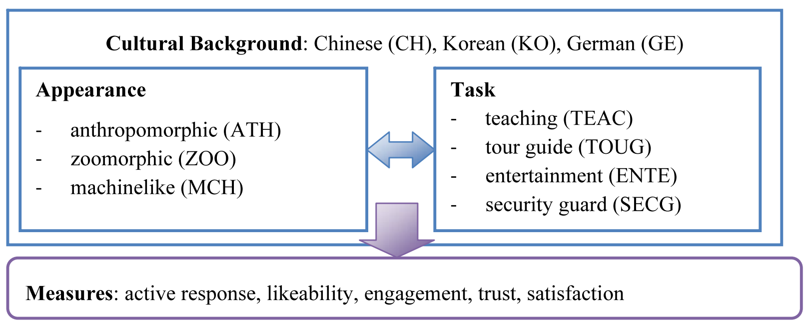

Fig. 2 Research Framework (a model of the cultural background, a robot's appearance and task)

teaching is of high sociability (e.g. a language teacher), tour guide (e.g. a tour guide in a museum) and entertainment are of medium sociability, and security guard (e.g. a security guard of a building) is of low sociability. The dependent variables were measured through a questionnaire.

Four hypotheses were formulated based on the conceptual model: the main effect of culture, robot appearance, robot task, and the interaction between a robot's appearance and its task.

3.1 Effect of National Cultures

Hofstede labeled "Big and fast are beautiful" as a masculine value and "Small and slow are beautiful" as a feminine value [8]. According to his dimension scores on national cultures, Korean culture is middle masculinity, while German culture is high masculinity [11]. Service robots, especially those used for social purposes, are generally medium and small sized and move relative slow compared with industrial robots. Thus, we hypothesize that Korean participants perceive a service robot as more likeable and they are more likely to respond to the robot versus German participants.

In addition, China, Korea and Germany have significantly different exposure to robots. According to the survey of World Robotics 2009, Germany is the largest market for industrial robots in Europe, where robots are widely used in the automotive, machinery, metal and food industry, whereas China and Korea have a relative low robot density in the general industry [32]. In Germany, service robots are generally made by small companies, including start-ups and spin-offs from research labs, whereas in Korea the service robots, especially entertainment robots and domestic assistants, are largely produced by big companies (e.g. Honda) [33]. Korea also listed robotics as one of the ten next-generation growth engines and promised a robot in every home by 2020 [34]. It is clear that Germany has a high exposure to industrial robots and Korea prefers service robots. As a result, we anticipate that Korean participants have a higher active response to and more positive attitudes toward a service robot than German participants do, and Korean participants engage more with the service robot. We define positive attitudes towards a service robot as "people consider the robot likeable, trustworthy and satisfactory".

H1.1: Compared with German participants, Korean participants have higher active response to, engage more with, and have more positive attitudes towards a service robot.

H1.2: Compared with German participants, Chinese participants have higher active response to, engage more with, and have more positive attitudes towards a service robot.

3.2 Effect of a Robot's Appearance

As shown in the second section, previous research argued that before reaching the reverse point in Mori's "Uncanny Valley", the increasing human-likeness results in greater familiarity and preference. Since the anthropomorphic robot used in our study was not extremely humanlike, and had exposed mechanical parts, we propose it would not reach the reverse point in the "Uncanny Valley".

H2.1: Participants have higher active response, engage more with, and have more positive attitudes towards an anthropomorphic robot versus a machinelike robot.

H2.2: Participants have higher active response, engage more with, and have more positive attitudes towards a zoomorphic robot versus a machinelike robot.

3.3 Effect of a Robot's Task

We classified teaching as a high sociability task, tour guide and entertainment as medium sociability tasks, and security guard as a low sociability task. Based on different levels of sociability, we hypothesized that when the robot conducts a high sociability task, it can stimulate higher active response and engagement than when it conducts a low sociability task.

H3.1: Participants have higher active response and engagement when the robot carries out a teaching task versus a security guard task.

H3.2: Participants have higher active response and engagement when the robot carries out a tour guide task versus a security guard task.

H3.3: Participants have higher active response and engagement when the robot carries out an entertainment task versus a security guard task.

3.4 Interaction Between a Robot's Appearance and Its Task

We found some evidence from a related study that when the task requires high sociability, a more humanlike robot is preferred; when the task requires little sociability, a machinelike robot is preferred; and for entertainment tasks, a zoomorphic robot is preferred [35]. Thus, we hypothesized that in tasks requiring high sociability, people would prefer a more human-like robot.

H4.1: For the anthropomorphic robot, participants have greater active response, likeability, engagement, trust and satisfaction when it conducts the teaching task than when it conducts the security guard task.

H4.2: For the anthropomorphic robot, participants have greater active response, likeability, engagement, trust and satisfaction when it conducts the tour guide task than when it conducts the security guard task.

H4.3: For the zoomorphic robot, participants have greater active response, likeability, engagement, trust and satisfaction when it conducts the entertainment task than when it conducts the security guard task.

H4.4: for machinelike robot, participants have higher active response and satisfaction when it conducts security guard task than when it conducts teaching task.

H4.5: for machinelike robot, participants have higher active response and satisfaction when it conducts security guard task than when it conducts tour guide task.

H4.6: for machinelike robot, participants have higher active response and satisfaction when it conducts security guard task than when it conducts entertainment task.

4 Methodology

We designed a $3 \times 3 \times 4$ experiment to test the hypotheses, with national cultures (Chinese, Korean and German) and robot appearance (anthropomorphic, zoomorphic and machine-like) as between-subject factors and robot task (teaching, tour guide, entertainment and security guard) as a within-subject factor. Each participant interacted with the robot in four scenarios, and in each scenario, they were asked to answer questions, speak certain words or perform certain activities according to the robot's instructions. We used Latin Square Design to rule out the influence of the sequence of tasks [36], that is the participants were divided into four groups, each group performed the tasks in a certain sequence and participants were randomly assigned to a group. A *Wizard of Oz* approach was used, in which the robot was operated by an experimenter sitting behind a screen, yet appearing to be operating autonomously. The robot was operated by the same experimenter across all conditions.

All written materials, instructions and robot speech were in Mandarin Chinese for the Chinese participants, Korean for the Korean participants and German for the German participants. We translated and back translated the materials to ensure the consistency in meanings across cultures.

4.1 Participants

We recruited 108 participants (36 Chinese, 36 Koreans and 36 Germans) from a university campus and randomly assigned them to a condition. The mean age of Chinese participants was 24.00 (SD = 1.85) years old and 97% were male. Korean participants were 23.28 (SD = 3.05) years old on average and 36% were male whereas German participants were 23.75 (SD = 0.87) years old on average and 89% were male. Due to the limited sampling pool on campus, the gender ratio was not balanced across cultures. We will discuss the gender influence in the final section. The first language of all Chinese participants was Chinese; for Korean participants it was Korean, and for German participants it was German.

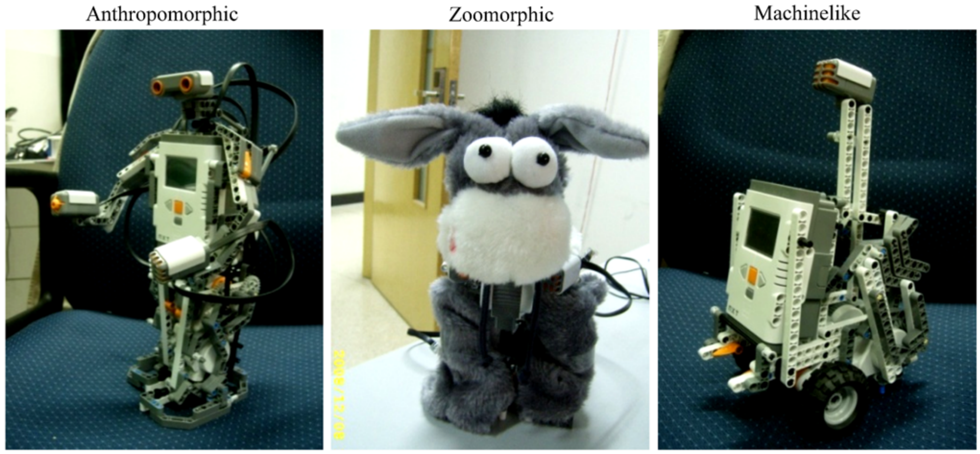

Fig. 3 Robot appearance in the experiment

4.2 Setup

We used the Lego Mindstorm NXT robot in our experiment. The experiment platform was developed using VBA in Microsoft Access and C# was used to connect with the robot and control its movements via a Bluetooth signal.

We designed three types of robot appearances (anthropomorphic, zoomorphic, machinelike), which were at different levels of human-likeness (as shown in Fig. 3). According to the coding schedule used to categorize the robot's appearance in one previous study, a robot with recognizable eyes and a humanoid body shape were more likely to be categorized as a humanlike robot whereas a robot with wheels and tracks and without facial features were generally categorized as machinelike robot [26]. In our study, to minimize the bias of participants' perception of the robot's gender and personality, we designed the robot to be as neutral as possible. For the anthropomorphic robot, we designed it with recognizable eyes, a head, arms and legs. For the zoomorphic robot, we covered the Lego parts with a donkey-shaped toy. For the machinelike robot, we eliminated any identifiable human body parts and assembled wheels to support the main part of the robot.

We used male voices in all conditions and set all voices at the same frequency (22 kHz 16 bit) and speed. Chinese voices were synthesized by Neospeech TTSApp (VW Wang); Korean voices were synchronized by the TTS engine at http://www.oddcast.com/home/tts, and German voices by AT&T Labs Text-to-Speech. The voices were broadcast from a speaker hidden at the rear of the robot and controlled in real time by the computer. We designed some motions for each robot appearance to facilitate interaction. For the anthropomorphic robot, the motions were shaking the head, lifting the arms and moving the legs; for the zoomorphic robot, the motion was nodding; for the machinelike robot, the motions were forward/ backward moving and turning.

Fig. 4 Experiment setup

We set up the experiment in the Usability Lab at Tsinghua University (see Fig. 4). The robot was set up on a desk, and the participant was asked to sit in a chair to interact with it. A camera was installed beside the chair, which was used to observe participants' response.

An experimenter sat behind a shielding screen to control the robot in a *Wizard of Oz* approach. The program for the robot's motions was stored in the control box on the robot's body. The sound files of the robot were stored in a computer. The computer and the control box were connected via a Bluetooth signal. When the experiment began, the experimenter observed the participant via the camera and controlled the robot's speech and motions to interact with the participant.

4.3 Task

We designed four tasks in this experiment (teaching, tour guide, entertainment and security guard). Participants were required by the robot to respond to its utterances or conduct certain actions. The content of the tasks were selected from common applications of each kind of service robot. In the teaching task, the robot acted first as an English teacher, asking the participant to finish an English test and repeat some English phrases, and then it acted as a physical education teacher and asked the participant to do some physical exercise. In the tour guide task, the robot introduced Copernicus's former residence and then asked the participant to answer some questions related to its introduction. In the entertainment task, the robot told two jokes, played music for the participant, and asked the participant whether to continue the interaction. The jokes were initially selected by a Chinese researcher and subsequently tested on two German students and one Korean student to assure the effect is comparable across cultures. In the security guard task, the robot introduced its detailed functions and checked the participant's student identity card. In each task, the robot initiated the interaction. We controlled the interaction in each task to be finished within the same time duration. To rule out the influence of task sequence, we randomly assigned the participants into four groups according to the Latin Square design and each group of participants took part in the four tasks in the same sequence.

4.4 Measures

The paper-based questionnaire after each scenario contained six parts: active response, likeability, engagement, trust, satisfaction and overall evaluation. We used the original English version of the scales to avoid mistranslation.

Active response: Active response evaluated participants' active response to social cues in human-robot interaction. For this feature, we adopted Lombard and Ditton's subscale (presence—active interpersonal), which measures one aspect of presence. (The adaptations in this case simply were substitution of "media environment" in the original scale with "human robot interaction" and substitution of "someone" with "robot"). The reported Cronbach's alpha was 0.77 [37].

Likeability: Likeability for the robot is a very important issue to smooth the interaction and strongly correlates with trust in the robot. We used Reysen, S.'s likeability scale, with a reported Cronbach's alpha of 0.90–0.91 [38].

Engagement: The engagement scale measures participants' involvement in the interactive experience. The scale was from Lombard and Ditton's scales for measuring six aspects of presence. (The adaptation in this case was simply substitution of "story" in the 6th item of original scale with "experience"). The reported Cronbach's alpha was 0.9 [37].

Trust: The trust scale evaluates whether the robot is believable and trustworthy as an interaction partner. Nicholson et al. defined trust as "confidence in the other party's reliability and integrity" [39]. To measure the trust in the robot, we used the SHAPE Automation Trust Index (SATI) trust scale; the only adaption here was substitution of "automation tool" with "robot" [40].

Satisfaction: The satisfaction scale evaluates whether the participants are satisfied with the interaction process. Satisfaction together with efficiency and effectiveness are the three aspects of the usability test. We measured satisfaction with the general satisfaction subscale of QUIS (questionnaire for user interface satisfaction); the reported Cronbach alpha was 0.94 [41]. The only adaption here was substitution of "the software" with "robot".

Overall evaluation: We asked participants to rate their overall perception of active response, likeability, engagement, trust and satisfaction according to the 7-point Likert scale (1-lowest, 7-highest) in order to get an overview of the above five features measured.

5 Results

The calculated internal consistency of the five scales is listed in Table 2. The internal consistency was higher than 0.7, which indicates acceptable reliability of the scales.

We conducted the One-Sample Kolmogorov-Smirnov Test to test the normality of the dependent variables. The data for likeability, engagement, trust and satisfaction passed the normality test, while the data for active response did not. Levene's Test indicated that for all variables the equality of error variances was assumed at the significant level of 0.05. For the five dependent variables, the correlation of participants' response in the four tasks was significant at the 0.01 level (2-tailed). Thus, we conducted Repeated Measure ANOVA to analyze the likeability, engagement, trust and satisfaction, and conducted the nonparametric test for several independent samples, Kruskal-Wallis H, to analyze the active response. The α for the model will be set at .05. Since the gender ratios across the three culture groups were not well balanced and we anticipate a participant's gender will influence his/her perception of and response to the robot, we added gender as a covariate in the statistic model.

5.1 Effect of Cultural Background

The hypotheses about the effect of cultural background were confirmed for all variables measured except active response.

Table 2 Internal consistency of the five measures

Feature	Items	Cronbach's alpha
Active response	3	0.721
Likeability	10	0.912
Engagement	6	0.874
Trust	7	0.841
Satisfaction	6	0.878

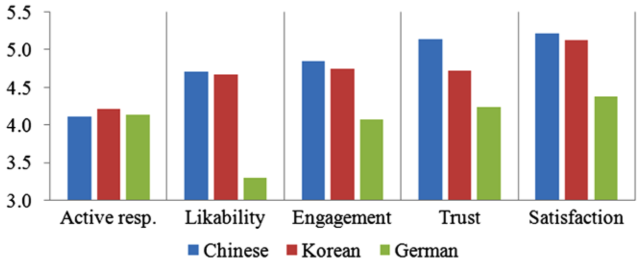

Fig. 5 Plot of main effect of cultural background for the five scales

Table 3 Post hoc analysis of main effect of culture

		CH	KO
Likeability	KO	0.784	
	GE	0.000*	0.000*
Engagement	KO	0.382	
	GE	0.000*	0.003*
Trust	KO	0.032*	
	GE	0.000*	0.042*
Satisfaction	KO	0.489	
	GE	0.000*	0.000*

The result showed that participants from China, Korea and Germany were significantly different on four scales ($p = 0.000$), namely, Likeability ($F(2,98) = 27.332$, Observed power $= 1.000$), Engagement ($F(2,98) = 8.595$, Observed power $= 0.964$), Trust ($F(2,98) = 10.110$, Observed power $= 0.984$) and Satisfaction ($F(2,98) = 11.847$, Observed power $= 0.994$). We did not find significance on the scale of active response (see Fig. 5).

The Post hoc analysis (LSD) showed that at the 95% confidence interval, German participants have significantly lower likeability, engagement, trust and satisfaction scores than Chinese and Korean participants, and Korean participants have significantly lower trust scores than Chinese participants do (see Table 3). Overall, German participants have the lowest rating in the four scales.

According to Hofstede's dimension of culture, Chinese and Korean culture are highly collectivist and have medium masculinity, while German culture is highly individualist and has high masculinity [42]. In the dimension of collectivism-individualism, previous studies showed that people from collectivist cultures are more likely to be influenced by people who communicate with them and tend to accept others' suggestions. This may explain why they had greater engagement in the interaction and higher trust in the robot. In addition, masculine cultures prefer large and quick robots to small and slow ones, while people from feminine cultures hold the opposite opinion. This may have caused Chinese and Korean participants to have higher likeability for and satisfaction with the small and slow social robot used in the experiment.

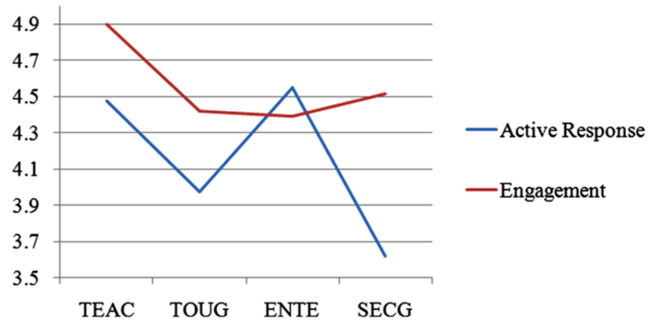

Fig. 6 Plot of main effect of robot's task on active response and engagement

5.2 Effect of Robot Appearance

Hypothesis 2 about the effect of the robot's appearance does not hold except on the scale of likeability ($F(2,98) = 4.624$, $p = 0.012$, Observed power $= 0.769$). Post hoc analysis on the likeability data indicated that there was a significant difference only between the zoomorphic robot and the machinelike robot ($\bar{X}_{\mathrm{ZOO}} = 4.588$, $\mathrm{SD}_{\mathrm{ZOO}} = 1.337$; $\bar{X}_{\mathrm{MCH}} = 3.880$, $\mathrm{SD}_{\mathrm{MCH}} = 1.121$; $p = 0.007$).

This result indicates that a robot's human- or animal-likeness will influence its likeability and the robot with the least humanlike appearance will receive the lowest likeability score, which confirms Syrdal's finding [22]. In addition, one previous study argued that zoomorphic robots are more suitable for social interaction with humans [35], which was also observed in our study.

5.3 Effect of Task

Hypothesis 3 is well supported when comparing the teaching task with the security guard task but partially fails when comparing the tour guide and entertainment tasks with the security guard task. The statistical analysis indicated that participants have significant difference in the active response ($\chi^2(3,105) = 44.246$, $p = 0.000$) but marginal difference in their engagement ($F(3,294) = 2.540$, $p = 0.057$) when comparing the four tasks. The paired t-test showed that compared with the security guard task, participants have higher active response in the teaching task ($t = 6.334$, $p = 0.000$), tour guide task ($t = 2.397$, $p = 0.018$) and entertainment task ($t = 6.683$, $p = 0.000$); LSD analysis indicated that participants have significantly higher engagement in the teaching task than in the other three tasks (see Fig. 6).

This finding indicates that participants had higher active response in the tasks with higher sociability (teaching, tour guide and entertainment) than in the task with low sociability (security guard) and participants had the highest engagement in the teaching task. It is also interesting to note that significantly higher active response and engagement do not necessarily lead to significantly higher likeability, trust and

Table 4 Correlation between culture/appearance/task and the dependent variables

	Culture	Appearance	Task
Act	0.008	–0.089	–0.153**
Like	–0.470**	–0.121*	–0.055
Eng	–0.288**	–0.085	–0.117*
Tru	–0.343**	–0.081	0.059
Sat	–0.323**	–0.070	0.028

* Correlation is significant at the 0.05 level (2-tailed)

** Correlation is significant at the 0.01 level (2-tailed)

satisfaction; the task rated significantly lower in the active response and engagement was not rated lower on the other scales.

5.4 Interaction of Robot Appearance and Task

According to the results from the Repeated Measure ANOVA, Hypothesis 4, about the interaction between appearance and task, was not proven true. Participants did not show significant preference or higher engagement when the robot's appearance "matches" its task as expected. That is to say, our hypothesis that "compared with low sociability tasks, a humanlike robot is preferred to conduct high sociability tasks, while a zoomorphic robot is especially favorable for entertainment tasks" is not supported. We will discuss the reasons in the general discussion.

Since there is no interaction between a robot's appearance and task, we conducted partial correlation analysis between appearance/task and the dependent variables with gender as a control variable. The result indicated that the robot's task has a closer correlation with the participants' active response and engagement than the robot's appearance does, while the appearance has a closer correlation with the likeability of the robot than the task does (see Table 4). This finding showed that the robot's appearance and its task might have influenced different aspects of participants' perception. The correlation analysis of culture and the dependent variables indicated that culture played a dominant role in determining a participant's trust in and satisfaction with the robot.

5.5 Interaction of Culture and Task

We found no significant interaction between the participant's culture and the robot's appearance, but the interaction of culture and task had a significant impact on engagement ($F[6, 294] = 2.591,\ p = 0.018$). The interaction effect showed that the Chinese and Korean participants had smaller differences in their engagement throughout the four tasks than the German participants did (see Fig. 7).

Table 5 Partial correlation of dependent variables with gender as controlling variable

	Act	Like	Eng	Tru
Like	0.434*			
Eng	0.487*	0.672*		
Tru	0.332*	0.700*	0.636*	
Sat	0.351*	0.719*	0.646*	0.774*

* Correlation is significant at the 0.01 level (2-tailed)

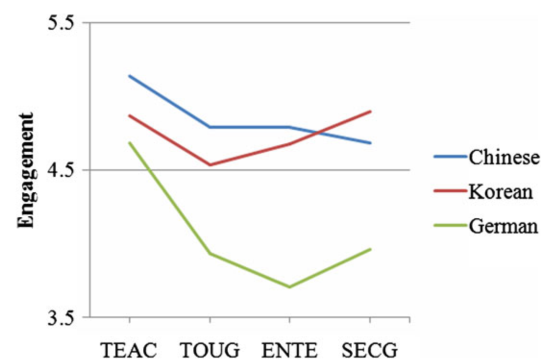

Fig. 7 Interaction of culture and task on engagement

This effect can be explained by the high-low context in Hall's dimensions of culture [12]. German culture is a typical a low-context culture, in which people depend highly on direct spoken information, while Chinese and Korean cultures are high-context cultures, in which people are influenced more by the context information and body language during the interaction. In this experiment, the tasks are mainly differentiated through the robot's utterances without variations in the context or the body language. Therefore, the engagement of Chinese and Korean participants was less influenced than that of German participants. This finding indicates that the low sociability task may reduce the engagement of people from low-context cultures more than that of people from high-context cultures.

5.6 Correlations Between the Five Measures

We analyzed the correlation between the five dependent variables, and the result indicates that all five measures had significant correlations at the 0.01 level (see Table 5). This finding agrees well with a previous study in the area of human-computer interaction [43], which indicated the positive correlation between the user's preference and interaction performance. It is clear that a highly likeable robot will engender high trust and satisfaction in the participants, and engagement has a strong positive correlation with the perception of the robot's likeability.

6 General Discussion and Conclusion

The results strongly suggest that cultural differences exist in participants' positive attitudes towards and engagement with a service robot. Compared with German participants, Chinese and Korean participants perceived the sociable robots used in the experiment to be more likeable, trustworthy and satisfactory, and they had higher engagement with the robot. This finding is in line with a previous study showing that German participants had more anxiety and more concern about robot's negative influence than Chinese participants did [44]. We propose that because of their masculine culture and individualism as well as high exposure to industrial robots, German participants may prefer to have more sense of control and to use robots as tools or machines (e.g., welding robots, spraying robots, robotic arms) rather than using them as companions or personal service providers. People may argue that the Korean group has more female participants that may influence the positive attitudes towards robots. We justify that among the three culture groups, the Chinese group has the largest proportion of male participants (97% male). If gender plays a more central role in predicting people's positive attitudes towards the social robots than cultural backgrounds do, we should anticipate that the Korea participants have the highest positive attitudes rather than the Chinese participants. Since we added gender as a covariate in the statistical model, the statistical power of testing the cultural effect was increased. We tend to conclude that the cultural background predicts people's positive attitudes towards the social robots. Countries with a high masculinity and high individualism culture or have a high exposure to industrial robots may have less positive attitudes towards the social robots.

Our hypotheses on the effect of robot appearance were not verified except on the scale of Likeability. Participants considered the zoomorphic robot to be more likeable than the machinelike robot. This result is reasonable since both the anthropomorphic and the machinelike robot used in the experiment have exposed mechanical parts whereas the zoomorphic robot has some amusing cartoonish features. However, we found no significant difference on the scale of likeability between the anthropomorphic and the zoomorphic robot. It should be noted that the anthropomorphic robot used in the experiment only differentiates itself from the machinelike robot with a recognizable body shape, eyes and a head. The result may indicate that the merely humanoid features are capable to increase people's familiarity to the robots and consequently result in an increased likeability.

Our manipulation of high and low sociability tasks are supported on the measures of active response and engagement. Participants have higher active response and engagement in a high sociability task (e.g., teaching) than in a low sociability task (e.g., security guard). Compared with active response, engagement correlates more with the perceived likeability, trust and satisfaction.

We found no evidence to support the importance of matching the robot's appearance with its task. This could be because the tasks were mainly differentiated by the robot's utterances without varying its motions. That may have resulted in a low sense of engagement or insensitivity to the mismatch between robot appearance and task. However, in the post-experiment interview, participants indicated their preference for the zoomorphic robot in the entertainment task. They also indicated that the humanlike robot was most suitable for the security guard task and secondarily suitable for the tour guide task, while the machinelike robot was most suitable for the security guard task, which agrees well with the findings from previous studies [24, 35]. Thus we conservatively suggest that zoomorphic robots might be suitable for entertainment tasks, while machinelike robots are more suitable for low sociability tasks (e.g., security guard); we plan to verify our hypotheses in future studies that involve more persuasive scenarios.

We observed significant interaction between culture and robot task and tend to conclude that when the tasks are mainly differentiated by the robot's utterances, people from a low-context culture are more influenced in the engagement scale than people from a high-context culture are. We propose that when the interaction involves more body language or changes in context, the task may exert greater influence on people from a high-context culture. The findings in this study indicate that when a robot is designed to be multifunctional, for example, to function as a teacher, an entertainer or a security guard at different time points, for users from low-context cultures (e.g., Germany), their engagement may be extremely reduced when the task is low in sociability (e.g., security guard). The designers may consider applying more interactive utterances to the robot if constant engagement of the users is desired. On the other hand, if engagement is a variable to be manipulated, changing the robot's body language or the context may be more influential on people from high-context cultures compared with merely changing a robot's utterances.

The strong positive correlations among the five variables suggest that to make a likeable robot and gain high trust and satisfaction, it is helpful for the robot designers to improve users' active response and their engagement in the interaction. Users might perceive a sociable robot that requires interactive response from them or makes them engaged more likeable, trustworthy and satisfactory. Robot designers can improve the engagement of the users through increasing the sociability of the robot task, specifically increasing the interaction time and frequency.

This tentative study is a beneficial attempt to test people's reaction to and perception of social robots through real interactions. We confirmed the cultural differences in people's positive attitudes towards and engagement with a social robot in different tasks. We suggest that to maintain similar positive attitudes and engagement with a social robot across cultures, the robotic designers should be aware of theses cultural differences and tailor the design to adapt to different cultures. In the future, a social robot that functions in a multi-cultural environment should be capable of detecting the users' cultural background and adapting its speech as well as motions to engage the users actively. Future study is expected to analyze how a social robot engages and maintains its engagement with users from different cultures.

Acknowledgements This study was funded by National Science Foundation China grant #70971074. We sincerely thank Hua Qin, Ting Zhang and Lin Wang of the Institute of Human Factors & Ergonomics for their help with many aspects of the experiment.

References

1. IFR Statistical Department (2009) Professional service robots are establishing themselves. In: World robotics 2009—service robots. IFR Statistical Department, Frankfurt
2. SuperDroid. Rp2w two way remote presence robot. Available from: http://www.robotshop.ca/superdroid-rp2w-remote-presence-robot.html
3. Sony. Aibo. Available from: http://support.sony-europe.com/aibo/index.asp
4. Takara. Tera bots. Available from: http://www.engadget.com/2005/01/20/takaras-new-tera-bots/
5. Aldebaran-Robotics. Nao. Available from: http://www.aldebaran-robotics.com/en/
6. Powers A, Kiesler S (2006) The advisor robot: tracing people's mental model from a robot's physical attributes. In: HRI'06, 2006
7. Kiesler S, Goetz J (2002) Mental models of robotic assistants. ACM, New York
8. Hofstede G (1980) Motivation, leadership, and organization: do american theories apply abroad? Organ Dyn 9(1):42–63
9. Hofstede G (1983) The cultural relativity of organizational practices and theories. J Int Bus Stud 14(2):75–89
10. Hofstede G, Hofstede G (2005) Cultures and organizations: software of the mind. McGraw-Hill, London
11. Hofstede G Cultural dimensions. Available from: http://www.geert-hofstede.com
12. Hall E (1977) Beyond culture. Anchor, Garden City
13. O'Neill-Brown P (1997) Setting the stage for the culturally adaptive agent. In: 1997 AAAI fall symposium. AAAI Press, Menlo Park
14. Reinecke K, Bernstein A (2007) Culturally adaptive software: moving beyond internationalization. Lect Not Comput Sci 4560:201
15. Bartneck C et al (2005) A cross-cultural study on attitudes towards robots. In: The 11th international conference on human-computer interaction (HCI '05). Las Vegas, USA
16. Bartneck C et al (2007) The influence of people's culture and prior experiences with Aibo on their attitude towards robots. Artif Intell Soc 21:217–230
17. Bartneck C (2008) Who like androids more: japanese or U.S. americans? In: 17th IEEE international symposium on robot and human interactive communication (RO-MAN'08)
18. Rau P, Li Y, Li D (2009) Effects of communication style and culture on ability to accept recommendations from robots. Comput Hum Behav 25(2): 587–595
19. Wang L et al (2010) When in Rome: the role of culture and context in adherence to robot recommendations. In: Proceeding of the 5th ACM/IEEE international conference on human-robot interaction. ACM, Osaka
20. Evers V et al (2008) Relational vs. group self-construal: untangling the role of national culture in HRI. In: 3rd ACM/IEEE international conference: human-robot interaction. ACM, Amsterdam
21. Fong T, Nourbakhsh I, Dautenhahn K (2003) A survey of socially interactive robots. Robot Auton Syst 42(3–4):143–166
22. Syrdal D et al (2007) Looking good? Appearance preferences and robot personality inferences at zero acquaintance. In: AAAI—spring symposium 2007, multidisciplinary collaboration for socially assistive robotics. AAAI Press, Menlo Park
23. Goetz J, Kiesler S, Powers A (2003) Matching robot appearance and behavior to tasks to improve human-robot cooperation. In: IEEE international workshop on robot and human interactive communication (RO-MAN '03)
24. Robins B et al (2004) Robots as assistive technology—does appearance matter? In: IEEE international workshop on robot and human communication (RO-MAN '04)
25. Kiesler S, Goetz J (2002) Mental models of robotic assistants. In: CHI '02 extended abstracts on human factors in computing systems. ACM, Minneapolis
26. Woods S (2006) Exploring the design space of robots: children's perspectives. Interac Comput 18(6):1390–1418
27. Mori M (1970) The uncanny valley. Energy 7(4):33–35
28. Gee F, Browne W, Kawamura K (2005) Uncanny valley revisited. In: 14th IEEE international workshop on robot and human interactive communication (RO-MAN '05). Nashville, USA
29. Walters M et al (2008) Avoiding the uncanny valley: robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion. Auton Robots 24(2):159–178
30. Walters ML et al (2009) Preferences and perceptions of robot appearance and embodiment in human-robot interaction trials. In: Artificial intelligence and simulation of behaviour (AISB'09) convention. Edinburgh, Scotland
31. Lohse M, Hegel F, Wrede B (2008) Domestic applications for social robots-an online survey on the influence of appearance and capabilities. J Phys Agents 2(2):21
32. IFR Statistical Department (2009) Executive summary of world robotics 2009, IFR Statistical Department
33. Kumar V, Bekey G, Zheng Y (2006) Industrial, personal and service robots. In: Bekey G (ed) Assessment of international research and development in robotics. World Technology Evaluation Center, Lancaster
34. Onishi N In a wired South Korea, robots will feel right at home. In: The New York Times, April 2 (2006)
35. Lohse M et al (2007) What can I do for you? Appearance and application of robots. In: Artificial intelligence and simulation of behaviour (AISB '07)
36. Montgomery D (1991) Design and analysis of experiments
37. Lombard M et al (2000) Measuring presence: a literature-based approach to the development of a standardized paper-and-pencil instrument. In: Third international workshop on presence. Citeseer, Delft
38. Reyen S (2005) Construction of a new scale: the Reysen likeability scale. Soc Behav Pers 33(2):201–208
39. Nicholson CY, Compeau LD, Sethi R (2001) The role of interpersonal liking in building trust in long-term channel relationships. Acad Mark Sci 29(1):3–15
40. Adams B et al (2003) Trust in automated systems. Ministry of National Defence

41. Chin J, Diehl V, Norman K (1988) Development of an instrument measuring user satisfaction of the human-computer interface. ACM, New York, pp 213–218
42. Hofstede G (1984) Culture's consequences: international differences in work-related values. Sage, Thousand Oaks
43. Britton C et al (2002) An empirical study of user preference and performance with UML diagrams. In: Proceedings of IEEE 2002 symposia on human centric computing languages and environments (HCC02). Arlington, Virginia
44. Bartneck C et al (2005) Cultural differences in attitudes towards robots. In: Robot companions: hard problems and open challenges in robot-human interaction (AISB '05). University of Hertfordshire, Hatfield

Dingjun Li received a joint Master's degree of Management Science & Engineering from Tsinghua University and Production System Engineering from RWTH Aachen in Germany in February 2009. Her research interests include cross-cultural design, human-robot interaction and human-computer interaction. She has worked at the State Intellectual Property Office of P.R. China since 2009.

P.L. Patrick Rau is the Director of the Institute of Human Factors & Ergonomics, and a Professor in the Department of Industrial Engineering at Tsinghua University in Beijing. He obtained his PhD from Purdue University in 1998. He has been a Visiting Scholar at Microsoft Research Asia in China, Visiting Professor at RWTH Aachen in Germany, and a Visiting Professor at Chuo University in Japan. His research interests include cross-cultural design, human-computer interaction, user experience, mobile computing, technology acceptance and driving safety. Prof. Rau is the chair of the China Board of the HCI International 2007, a member of the Human Factors and Ergonomics Society, the Association for Computing Machinery, China Ergonomics Society, and Taiwan Ergonomics Society.

Ye Li is a Master Candidate at the Institute of Human Factors & Ergonomics at Tsinghua University. She completed her graduate study at the Laboratory for Machine Tools and Production Engineering at RWTH Aachen in Germany in March 2009. She received her Bachelor's Degree from the Department of Industrial Engineering at Tsinghua University in 2007. Her current research interests include cross-cultural design, human-robot interaction and technology acceptance.