Abstract—Recent developments in humanoid robotics have made possible technologically advanced robots and a vision for their everyday use as assistants in the home and workplace. Nonetheless, little is known about how we should design interactions with humanoid robots. In this paper, we argue that adaptation for user attributes (in particular gender) and task structure (in particular a competitive vs. a cooperative structure) are key design elements. We experimentally demonstrate how these two elements affect the user’s social perceptions of ASIMO after playing an interactive video game with him.

I. INTRODUCTION

Honda’s ASIMO humanoid robot (Figure 1) is intended to be a first step toward a long term vision of an intelligent system that would “duplicate the complexities of human motion and actually help people” [1]. Humanoid robots might provide day-to-day support in the home and the workplace, doing laundry or dishes, assisting in the care of the elderly, or acting as a caretaker for individuals within a home or institution. Many of these tasks will involve a close interaction between the robot and the people it serves.

This vision of humanoid robots is far-reaching with implications for robot design, perception, manipulation, and interaction design. Researchers have studied the humanoid robot itself, including attributes of personality [2], [3], appearance [3], [4], and social and emotional expressiveness [5]–[7]. The design of the interaction between the user and the robot has been less well studied although peer-to-peer human-robot collaboration [8] and issues of user perspective [9] had been explored.

Figure 2 presents diagrammatically the critical “social” aspects of the design of a human-robot interaction: robot attributes (e.g. appearance, character), user attributes (e.g. gender, age), and task structure (e.g. cooperative, hierarchical). In this paper, we explore some of the dimensions of user attributes and task structure, and experimentally demonstrate their effects with Honda’s humanoid robot, ASIMO.

Nass and his colleagues have demonstrated systematically that user attributes such as age, gender, and personality type are critical aspects of interactive computer interfaces [10], [11]. Woods et al. extended some of their results to mobile robots and showed that such attributes affect how users ascribe their own personalities to a robot’s personality [12]. We believe that these attributes will play an even greater role in the design of interactions with a humanoid robot because the robot has a human form.

Elements of task structure also play a critical role in the interaction design of a humanoid robot. A robot’s task requires that it have certain skills, knowledge, resources, and behaviors. These attributes will be represented in the interface and interactions of the robot. However, only a few studies have considered the nature of the robot’s task or tasks as part of the design [4], [13], [14].

The study presented in this paper explores how user attributes and task structure affect responses to ASIMO. We investigated the user attribute of gender and task structures of competition and cooperation. We explored these questions with two experiments: a human-human experiment and a human-robot experiment. Participants interacted with a second human participant or with ASIMO while playing a video game. We manipulated whether the video game was cooperative or competitive. Men and women showed significant differences in their social experiences and perceptions of ASIMO based on the task structure. Our findings suggest that, to be successful, the interaction style of humanoid robots will need to be customizable to the attributes of the people and to the task at hand.

II. BACKGROUND

Social robots such as museum tour guides, assistants for the elderly, tutors for children, and receptionists [6], [15]–[18] must be designed to set appropriate expectations. Many of the people who will interact with these robots will have
no training or skills in how to do so. How these robots should adapt to their users and the tasks at hand has not yet been extensively explored. We suggest that such exploration can be heavily informed by research in human-computer interaction, which has systematically examined the tendency to respond to computers in a social and natural way. Some researchers report that people recognize personality traits in a computer, through relatively simple social cues and social interaction [20], [21]. Nass and his associates have shown that people prefer computer personalities that match their own personality type, and to evoke more cooperative behavior [22].

These studies imply that social responses to interactive computers are similar to the way that people respond to other people [21], [23]. For example, when people used an interactive computer interface in a collaborative group task, their behavior was similar to their behavior when they collaborated with a person on the same task. Nonetheless, in experiments that have a human or comparison group, participants do not always find the computer to be identical to the person [24], [25]. In one study, the computer was perceived to be friendlier, more trustworthy, to resemble their own personality type, and to evoke more cooperative behavior [20].

Will embodied robots and humanoid robots in particular trigger the same responses as software agents and information and communication systems? Although the research suggests that they might, this claim needs to be tested. The social science literature provides a framework for hypothesizing how individual attributes such as gender will affect human robot interaction. It also suggests how a robot’s task structure may affect people’s mental representations of their social interaction partners and their experience with the robot. Knowledge from the literature on individual attributes and task structure helped form the premises for our study.

I. THEORETICAL BACKGROUND

The structure of the task has an impact on the behavior of individuals and groups. To assess these differences, McGrath developed a typology of tasks for groups [19]. The model consists of two bipolar dimensions, ranging from competitive to cooperative tasks and conceptual to behavioral. Our research focuses on the competitive to cooperative dimension. In laboratory studies of competitive situations, the success of one group member causes the failure of another [26]. In cooperative situations, one group member’s success increases the chances of success for the entire group. Research has shown that members of cooperative groups display more friendliness and encouragement towards other members of the group, and evaluate their group members more positively, whereas members of competitive groups depend less on each other, and display less friendliness and more rejection towards group members [27], [28]. It is possible that when people interact with ASIMO in a competitive or a cooperative task structure, they might show similar tendencies in their perceptions of ASIMO and in their own feelings.

Individual attributes also shape social experiences. Current research in HRI has not taken this aspect of interaction design into account in depth. Kanda and his associates found in their experiments with Robovie that a robot could develop a more successful relationship with children if it paid attention to individual differences such as English proficiency [16].

Our study examines the impact of the user’s gender, a human attribute well-known to have a significant impact on social behavior [29]. Social role theory suggests that gender differences in social behavior arise from the social roles held by men and women [30]–[32]. Studies have shown that in both individual and group contexts, men generally display behavior that is more dominant, aggressive, and task-oriented, whereas women tend to be more socially oriented, sensitive, warm, and caring [30]. Women also have been found to be more able to decode nonverbal cues in a way that enhances their social skills [31]. These gender-based differences suggest that men’s and women’s perceptions of ASIMO might be different. Women may care more about a robot’s social skills; men may care more about its task skills.

Gender has also been shown to have an effect on social interaction over a variety of tasks. Common gender stereotypes suggest that men are more competitive than women [27]. Men develop tactics to maximize their benefit, while women are less concerned with the task, and more oriented to the social setting [33]. These results suggest that women and men might show differences in their perceptions of ASIMO and their affective state for both competitive and cooperative tasks. We suggest that men will respond more to changes in task structure because they are attuned to differences in the task behavior of the robot.

Hypotheses

Drawing ideas from these findings in the social science literature, we formulated two hypotheses about responses to interactions with ASIMO. The hypotheses are related to the user attribute of gender and the task structure of competition and cooperation.

• Hypothesis 1. People who interact with a robot in a cooperative task will perceive the robot more positively than people who interact with the robot in a competitive task.
• Hypothesis 2. The user’s gender will shape social responses to robots in cooperative and competitive tasks. Women’s perceptions will be based on social attributes of the robot and will not change based on the task structure, while men’s perceptions will be shaped by the task.

II. METHOD

We designed an interactive videogame for a person and ASIMO and then created two experiments around that experience. In the first experiment, two human participants played the video game against or with each other; in the second experiment, one human participant played the video game against or with ASIMO. The experiments used a between-subjects design.

A. Interaction Design

The interactive experience was designed with the goal of having ASIMO and the participant act as peers. We intended for the experience to be engaging and to provide sufficient opportunity for interaction even given limited sensing of the participant. We implemented a two-player video game that is similar to the Mirror Time game introduced by Sony for their EyeToy system [34]. The game provides ASIMO and the human participant with identical screens and allows them to interact along a diagonal line of sight (Figure 3).

The screen that the participant sees shows a video (mirror) image of him or herself and four targets colored red or green. The goal of the game is for the participant or ASIMO to swipe his/her hands over the green targets while avoiding the red targets. Green targets increase the score, red targets decrease it. In the starting configuration, the task is easy because of the straightforward representation of the participant in the image. As the game progresses, however, the participant's view of him or herself is flipped horizontally/vertically and/or split into two/four (Figure 4).

ASIMO plays the game well because the flipping of the image is relayed to his controller and does not confuse his motions. The human participants, on the other hand, are often confused by the various permutations of the image and their reaction time and errors increase. We scaled back the opportunity for interaction even given limited sensing of the participant.

ASIMO interacts with the participant at three points during the experience: introduction, game, and conclusion (Figure 5). The introduction and conclusion are pre-recorded speeches directed at where the participant is standing (as directed by the experimenter). The speeches were recorded by a professional female actor using the voice of a child as ASIMO’s size and public presentation suggests the attributes of a 10 year-old boy. During the game, ASIMO interjects encouraging comments as he and the participant play. For example after a missed point, he might say, “Oh, don’t worry about it.” Or after a good score, he might say “Good job!” All of the comments are positive so that they can be used in both the competitive and cooperative conditions.

Fig. 4. Images were flipped horizontally (b) and vertically and split into two (c) or four (d) regions that flipped separately.

ASIMO’s behavior was controlled by an adaptive algorithm and supplemented with Wizard-of-Oz techniques. The algorithm monitored both ASIMO’s and the participant’s performances and adapted ASIMO’s behavior to give appropriate feedback to the user. The Wizard-of-Oz techniques were used to locate the user so that ASIMO could gaze at the user and initiate/end the interaction.

B. Experimental Design

We conducted two between-subjects experiments. In the first experiment, participants interacted with a male college-age human partner. In the second experiment, participants interacted with ASIMO. We manipulated the video game to be either competitive or cooperative. In the competitive structure, participants competed against their partner to maximize individual success and both their scores were shown on the screen. The cooperative structure asked participants to work with their partner towards a given goal that was shown on the screen along with the group score.

In each experiment, participants were first given a brief description of the overall experiment procedure. In the human-human experiment, participants were recruited such that the two players would not know each other and they were introduced to one another by the experimenter. In the human-robot experiment, ASIMO introduced himself. After the introduction, participants were asked to answer a pre-experiment questionnaire regarding their affective state and their impressions of their partner. The questionnaire was followed by providing participants with the details of their task. The experiment manipulation was introduced at this point. Half of the participants were told that they were going to compete with their partner, while half were told that they needed to work together with their partner towards a given goal. They performed a short practice round of the experimental task followed by four full length rounds of increasing difficulty. In each round, their compensation increased if they beat their partner in the competitive structure or if they met the given goal in the cooperative conditions.

Fig. 3. Physical orientation of the game.
structure. After completing the task, they answered a post-experiment questionnaire regarding their affective state, their perceptions of their partner, and their demographic information. During the experiment, two experimenters, one male and one female, were present in the room.

C. Measures and Sample

The task as well as the experimental procedure was identical across the competitive and cooperative manipulation except for the instructions and the presentation of the score. We controlled all factors except the two independent variables: the competitive vs. cooperative task structure (a manipulated independent variable) and the participant’s gender (a measured independent variable). The dependent variables measured were the participant’s own affective state, their perception of their partner’s affective state, how connected they felt to their partner, their perceptions of their partner’s physical, social, and intellectual characteristics, and participant demographics. To measure these variables, we used scales that were previously tested and proven to be reliable in evaluating people (Table 1). Scales of affect and interpersonal connectedness were used both before and after the experimental task in order to subtract pre-experiment biases and capture how variables changed due to participants’ involvement in the task.

We recruited 24 (6 females, 18 males; average age 21) and 26 (16 females, 10 males; average age 21) Carnegie-Mellon undergraduate students, diverse in undergraduate major, for the human-human and human-robot experiment respectively. Participants were paid a base rate of $10. They received a bonus of $1 for each round of the game that they won over the other player in the competitive game or, in the cooperative game, if they met the given goal with their partner. Since ASIMO was most often perceived as a male gendered robot due to its physical design and boy-like voice, we only used data with male partners in the human-human experiment, requiring us to drop data from 6 participants. We averaged the ratings of players when they were both male due to the interdependency of data from two players, which left us with 12 (6 females, 6 males) data points for the human-human experiment. All 26 data points were used in the human-robot experiment.

III. RESULTS

Our data analysis used three methods; repeated measures analysis of variance (MANOVA), regression (Least Squares Estimation), and multivariate correlations. The first method applied an Omnibus F-Test to see if the difference between pre-experiment and post-experiment measurements was significant across the two experiments, task structures, and/or genders. The second technique used a linear regression on the variables that were significant across conditions to identify the direction of main effects and interactions. The last method looked at how these variables correlated with each other. We also ran reliability tests and factor analyses on the scales we used for measurement.

We compared item reliabilities for all partner evaluation measures across our two experiments. One item in the sociability scale (five items) was removed to attain a reliable scale in both our experiments. The attractiveness scale (two items) showed high reliability in our human-human experiment as well as in the humanlike interface agent evaluation study, but the reliability in our human-robot experiment was significantly lower ($\alpha=0.05$). A similar pattern was observed in the mutual liking scale (two items) where the item reliability for our human-robot experiment was not only significantly lower but also negative ($\alpha=-0.16$). This result shows that some of the scales that were previously used to evaluate humans were not reliable in evaluating ASIMO in this interactive experience. Therefore, we did a factor analysis of all the partner evaluation items to identify scales that were reliable for evaluating both humans and robots. Our analysis identified two prominent factors, a five-item scale of desirability ($\alpha=0.80$) and a five-item scale of trustworthiness ($\alpha=0.85$).

A regression on desirability and trustworthiness scales showed no significant difference across experiments and task structures. The friendliness, dominance, pleasure, and arousal ratings of partners also did not show significant differences across experiments and task structures. In the human-robot experiment, involvement in the interaction was higher in the competitive task than in the cooperative task.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Evaluation</th>
<th>Questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive &amp; Negative Affect [35], [36]</td>
<td>Self</td>
<td>Pre + Post</td>
</tr>
<tr>
<td>Task enjoyment and difficulty [37]</td>
<td>Self</td>
<td>Post</td>
</tr>
<tr>
<td>Affect Grid Scale [38]</td>
<td>Partner</td>
<td>Pre + Post</td>
</tr>
<tr>
<td>Interpersonal Transaction Cycle [39]</td>
<td>Partner</td>
<td>Pre + Post</td>
</tr>
<tr>
<td>Inclusion of the Self in the Other [40]</td>
<td>Partner</td>
<td>Pre + Post</td>
</tr>
<tr>
<td>Social &amp; Intellectual Evaluation [41].</td>
<td>Partner</td>
<td>Pre + Post</td>
</tr>
<tr>
<td>Humanlike, attractiveness, mutual liking, trustworthiness [37]</td>
<td>Partner</td>
<td>Post</td>
</tr>
</tbody>
</table>

* A copy of the questionnaires and a video clip of the game can be found online at http://www.peopleandrobots.org/taskstructure.html.
When we included participant's gender in our statistical model, significant interactions appeared in several variables. Men and women varied greatly in their evaluation of their affective state and their partner's characteristics across task structures. Women did not vary in how desirable they rated ASIMO across conditions ($F[1:22]=0.28, p=0.60$), whereas men rated the robot more desirable in the cooperative task than in the competitive task ($F[1:22]=5.57, p=0.03$). This effect did not appear in ratings of a human partner.

Gender and task structure had strong interactions over how participants rated their own affective state. In the human-robot experiment, women did not show a significant difference in how they reported their positive affect ($F[1:22]=0.04, p=0.84$), as well as their levels of pleasure, involvement, and excitement (which are subscales of the positive affect scale) across the two task structures. Men reported their positive affect significantly higher in the competitive task than in the cooperative task ($F[1:22]=4.79, p=0.04$), which seems to be driven by a similar tendency in their involvement in the task ($F[1:22]=5.57, p=0.03$). This effect did not appear in ratings of a human partner.

Gender and task structure had strong interactions over how participants rated their own affective state. In the human-robot experiment, women did not show a significant difference in how they reported their positive affect ($F[1:22]=0.04, p=0.84$), as well as their levels of pleasure, involvement, and excitement (which are subscales of the positive affect scale) across the two task structures. Men reported their positive affect significantly higher in the competitive task than in the cooperative task ($F[1:22]=4.79, p=0.04$), which seems to be driven by a similar tendency in their involvement in the task ($F[1:22]=9.80, p<0.01$). On the other hand, their reports of their positive affect showed an opposite tendency in the human-human experiment where they reported their excitement significantly higher in the competitive task than in the cooperative task ($F[1:8]=5.17, p=0.05$).

Our results showed that people do differ in their perception of the robot based on task structure as our second hypothesis suggested. However, this effect was much stronger for men than for women, providing stronger support for our second hypothesis. The social experiences of our female participants (evidenced by similar ratings of their positive affect and their perceptions of the robot in the two conditions) did not rely on the structure of the task whereas the experience of the male participants was more dependent on the task. This implies that women might have relied on the social attributes of the robot (which were identical across the two conditions) when they evaluated ASIMO. However, more research needs to test this conclusion.

IV. DISCUSSION

Our experimental results point to a few initial speculations for the design of human-robot interactions. People's gender and task structure do affect their social experience with a robot. The results suggest that designers of interactive experiences should make sure that the interaction style of the robot fits the task structure and the individual attributes of users. For example, men found ASIMO less desirable in the competitive task than in the cooperative task, implying that ASIMO should act cooperatively with men when social desirability or acceptance is required. On the other hand, special attention should be paid to providing ASIMO with an appropriate set of social skills and non-verbal gestures for tasks that he will be performing with women.

The ability to adapt to users is particularly important for robots such as ASIMO that are expected to interact with men and women of different ages and backgrounds. This adaptation could be achieved either in real time by adjusting parameters of a single interface or through switching between multiple static interface designs.

Similarly, some robots, like ASIMO, are intended to handle a number of quite different tasks. We explored one dichotomy in the task structure, competitive vs. cooperative, but many others remain to be explored. Each of these will likely lead to design principles for successful interactive experiences and like our results, may well depend on the individual attributes of the user.

Our study had a number of limitations. First, ASIMO's interactivity is limited because of the difficulty of building sensing and natural language processing for interactive systems. We used Wizard-of-Oz techniques such as having ASIMO direct his gaze to the location where the participant was standing or speaking at a set time in the experiment, but these simple techniques are limited in the level of interactivity that they can produce. This problem was exacerbated by the engrossing nature of the video game, which permitted the participants to focus less on the interaction with ASIMO. In a future setup, a game that requires the user to imitate ASIMO's behavior, such as imitating his dance steps may minimize this limitation.

Task structure did not have a significant effect on people's perception of their partner. Additionally, men and women did not show differences in their perceptions across task structures in our human-human experiment. This discrepancy, inconsistent with findings in the social science literature, might have occurred because most studies use...
tasks such as negotiation that are more interactive than our experimental task with ASIMO.

Another limitation in our study was our participant sample. Some of the statistical analyses in our human-human experiment would have benefited from more participants. Furthermore, we had high variance in our data presumably because of the novelty effect and people’s inconsistent perceptions of robots in general.

This study is an early exploration in a fairly new research domain. We explored only a few dimensions of the factors that shape social interaction. Future experiments will focus on the development of additional principles for the design of socially appropriate human-robot “interaction”.

REFERENCES