Effects of Head Movement on Perceptions of Humanoid Robot Behavior

Emily Wang Constantine Lignos Ashish Vatsal Brian Scassellati

Computer Science Department
Yale University
51 Prospect St.
New Haven, CT 06510

emily.wang@yale.edu constantine.lignos@yale.edu ashish.vatsal@aya.yale.edu scaz@cs.yale.edu

ABSTRACT

This paper examines human perceptions of humanoid robot behavior, specifically how perception is affected by variations in head tracking behavior under constant gestural behavior. Subjects were invited to the lab to "play with Nico," an upper-torso humanoid robot. The follow-up survey asked subjects to rate and write about the experience. A coding scheme originally created to gauge human intentionality was applied to written responses to measure the level of intentionality that subjects perceived in the robot. Subjects were presented with one of four variations of head movement: a motionless head, a smooth tracking head, a tracking head without smoothed movements, and an avoidance behavior, while a pre-scripted wave and beckon sequence was carried out in all cases. Surprisingly, subjects rated the interaction as most enjoyable and Nico as possessing more intentionality when avoidance and unsmooth tracking were used. These data suggest that naïve users of robots may prefer caricatured and exaggerated behaviors to more natural ones. Also, correlations between ratings across modes suggest that simple features of robot behavior reliably evoke notable changes in many perception scales.

Categories and Subject Descriptors

I.2.9 [**Artificial Intelligence**]: Robotics – *Operator interfaces.*

General Terms

Human Factors, Experimentation.

Keywords

Head tracking behavior, Intentionality, Coding scheme.

HRI'06, March 2–4, 2006, Salt Lake City, Utah, USA.

1. INTRODUCTION

Psychology and sociology have established that perception of behavior occurs at various levels, some of which are subconscious. Perception of any behavior produced by a robot depends especially on the extent to which the robot is viewed as an anthropomorphic individual. Kiesler and Goetz [7] have applied Big Five personality traits to machines to measure anthropomorphic attributions, as well as more mechanistic measures. Their work suggests that people can harbor mental models of machines that call on human traits to characterize machine behavior, even though it is well understood by most observers that machines such as robots are strictly inanimate.

For characterizing a robot intended for naïve users, the amount of autonomy that the robot exhibits will, in part, define the nature of the robot's social role in relation to the user. Autonomous action is perceived as the intentional behavior of a living being [1], and it is very easily recognized, even when other features of the agent clearly indicate that it is not living. In a real-time, spontaneous social interaction, human individuals could be more likely to view robots as anthropomorphic. According to one theory,

> Interaction consumes cognitive resources. Our attention is a limited resource. When we are interacting with an animal or object, we can't think intellectually about what is really going on, for instance, a biological or computer program. Instead we focus on what the animal or object is doing and automatically make attributions as we do with people [6].

It may also be more comfortable for naïve users to interact with an unfamiliar robot when the situation seems social, since there are pre-existing mental scripts for navigating various social situations. DiSalvo and Gemperle argue that "anthropomorphic form is more than just an embellishment—anthropomorphic form can be understood and practiced as a means of solving design problems," [4] especially when used to make a product more familiar and to project human values onto a product. Since anthropomorphic perceptions of a robot product strongly influence a user's liking and trust for the robot, the ability to mediate these perceptions should lead to robot designs that make a more intuitive interface, which would facilitate use of the robots in the manner that is intended. Although more human-like behavior and likeness should improve user perceptions, Mori's "uncanny valley" suggests non-linearity in this effect [9]. At some point on the continuum of increasing human likeness, slight inaccuracies in the portrayal of a human being result in character that users find disturbing.

The following study investigates how variations in movement behavior can influence human perception of a robot from a short interaction, given an established structural anthropomorphic form. Structural form is the physical shape and configuration of an object. DiSalvo, Gemperle, and Forlizzi [5] also identify gestural behavior, personality, and demonstration of awareness as features of anthropomorphic form. These latter three features are complex products that arise from a robot's behavior.

The data analysis for this experiment also makes use of a coding scheme originally applied to narrative explanations of human behavior. Malle's F.EX coding scheme [8] was adapted for this study to measure subject's impressions of intentionality. In order for a robot to establish a social relationship with a human, "the human must believe that the robot has beliefs, desires, and intentions" [2].

We hypothesize that both obvious and non-obvious differences in motion behavior will profoundly influence human perception of a robot in a social setting. To test this theory, we have implemented four different head tracking modes on a humanoid robot, Nico. Nico does not have a particularly human face, but Nico's motor joints were designed to mimic the kinematics of a real human child. Only general arm gestures have been possible, since Nico does not have a hand with any DOF. While Nico lacks some finishing touches such as an attractive shell to cover the motor mechanisms, this system serves as a good basis for developing realistic human movement and observing subsequent effects on social interaction. Precisely because Nico's wires and motors are exposed, the body form is at first somewhat unrecognizable, and this allowed us to investigate the level of behavior that could prompt anthropomorphic identification of a robot.

2. METHODOLOGY

2.1 System Behavior

Nico is an upper-torso humanoid robot with body dimensions and kinematics constructed to match the 50-percentile male one-year-old human child. Nico has a seven-DOF head, a six-DOF right-arm and a one-DOF waist. For this research, two DOF were used in the head and four in the arm, as well as a short focal length camera mounted on the head. Instead of a functional hand, Nico currently has a 60mm steel shaft with a wooden ball on the end, which is attached to the wrist plate at a 90-degree angle.

A control program activated both the scripted arm gestures and the tracking program, and allowed the experimenters to restart the demonstration and select tracking modes for each trial. The tracking program relied on skin detection to locate the object most likely to correspond to a person's head. The camera image was fed through a threshold skin filter and then smoothed. Labeling the connected components of skin and choosing the highest region in the field of view above an additional size threshold yielded a reliable identification of the human being closest to Nico.

The position of the nearest human face was averaged over three frames as a compromise between accuracy and reaction time. To make up for a lag that would become apparent in tracking moving agents, the head motors received instructions to move double the distance that the face had traveled since the last iteration, under the assumption that the face would continue to move while Nico's head was in the process of orienting to the new location. A drawback of this built-in overshoot was that it tended to cause Nico's head to oscillate if the subject stood fairly still. This lead to the development of a suppression option that turned out to be a useful experimental manipulation. With suppression turned on, small, back-and-forth movements of the head motors were censored, so the head would converge immediately on a position and become still.

The tracking system paused after sending new (x, y) pixel offset coordinates to the head motors because the data from a camera in motion would have been noisy and constant monitoring was not necessary to produce realistic tracking.

Upon receiving a start message from the control program, the head would rise and enter the chosen tracking mode. After a delay of a few seconds, the arm would rise and wave to nobody in particular, and then return to the natural hanging position. There was then a pause in the gesture script, after which the arm would rise again and beckon for the hypothetical human agent to come closer. The arm then returned to the default hanging position, and stayed still for the remainder of the trial. Coordinate values and the rate of change used to script motor movements were determined by trial and error, and delays normally present in the motor code were removed to make the movements more natural. These wave-and-beckon motions would take place whether or not a human face had been detected, although this lack of contingency was not obvious to casual observers, including our subjects.

2.2 Experimental Procedure

Potential subjects were asked to come to the lab for a short, 1-minute interaction with Nico. They were told that we needed them to play with Nico in order to test a gesture-based method of communication that was still in the early stages of development. Participants were briefed with a written sheet, which stated that Nico could react to tone of voice, body movement, and basic hand gestures, but subjects were not otherwise instructed on specific strategies for interacting with the robot. Each subject was escorted into the closed laboratory, where only one experimenter was present to monitor the system and videotape the trials.

Directly after their interaction experience, subjects were asked to complete three surveys. A separate control group did not visit the laboratory for the interaction, but filled out the first two surveys from home. Experimental subjects were debriefed the next day with a statement revealing that Nico's gestures were completely scripted, and that the robot's head motion was based only on processing face position.

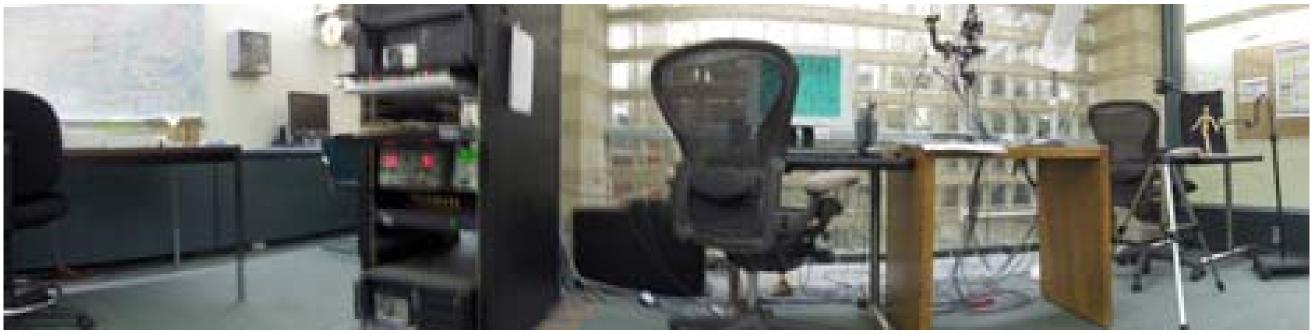

Figure 1. Panoramic view of the robotics laboratory used in Survey 1.

2.2.1 Conditions

Four head-tracking modes programmed for Nico corresponded to the four experiment conditions. In condition 1 (No Tracking), Nico's head was raised at the beginning of the trial, and then stayed still facing straight ahead. In condition 2 (Natural Tracking with Suppression), Nico's head changed position to face the largest piece of skin in sight, effectively following any human agent that moved slowly enough. In condition 3 (Tracking without Suppression), Nico continued to orient towards areas where skin was detected, but small oscillatory movements of the head that arose from repeated tracking adjustments were allowed to manifest themselves. In condition 4 (Avoidance without Suppression), Nico's head turned away from the largest detected region of skin.

2.3 Data Collection

Three different online surveys were created; the first two were administered to both the control and test groups, while the last survey was only given to experimental subjects who completed the laboratory interaction. Subjects were not allowed to navigate backwards through the pages of a survey.

2.3.1 Survey 1: Form and Body Identification

The questions on this survey used static photographs of Nico and the lab environment to measure the extent to which both groups recognized Nico's torso as structurally anthropomorphic. Respondents read the following statement that described anthropomorphic form:

> "Anthropomorphic form" is any imitation of the human form in non-human objects. It is often used in product design to make objects familiar, and to project human values.

The description and an accompanying illustration giving examples were adapted from an informational website describing work in anthropomorphism that is geared towards educated readers but assumes no specialty in any field of research [6]. Subjects were then presented with a panoramic scene of the robotics lab that included Nico, the server rack and workstation for Nico, and several objects posed in the photograph that were meant to be perceived as anthropomorphic forms (see Figure 1). Subjects were asked to list by name as many anthropomorphic forms as they could identify in the picture.

On the next screen, subjects were presented with a rotated picture of Nico (see Figure 2), and on the following screen, the same photograph upright. Nico is pictured in the default position, with all motors off so that the arm hangs straight down and the head is tilted forward and down. In a legend positioned to the side of the main photograph, a smaller, faded version of the photograph is split into 12 tiles of roughly equal size, labeled 1 to 12, and respondents were asked to select the number of the region in the photograph that contains particular examples of Nico's body parts (e.g. arm, hand, neck, mouth, eyes, and chest). Respondents were also given the choice to answer with "Nico does not have this body part," which would be the correct answer in regards to Nico's mouth.

2.3.2 Survey 2: Social Group Markers

The second survey asked respondents to identify Nico's age (e.g. Infant, Toddler, Child, Teenager, or Adult), gender (e.g. Male, Female, or Neither), and apparent Intelligence (e.g. Very, Better than expected, Average, A little, Not at all, No opinion). For the intelligence rating, the question was worded very specifically, "If you were in the same room as Nico, how smart do you think the robot would seem?" The purpose of this wording was to avoid triggering subjects' beliefs as to whether machines can be said to have intelligence.

Subjects were also invited to write a couple sentences explaining their assessment of Nico's age, gender, and intelligence. The last question in this survey recorded the subject's own age, gender and college major.

2.3.3 Survey 3: Perceptions of Behavior

This last survey used a 5-point scale (e.g. 5=Very much, 4=Somewhat, 3=Average, 2=A little, 1=Not at all, and blank=No opinion) and asked subjects to rate Nico on eight different attributes that were chosen to qualify their impression of the social play experience. To measure the extent to which they

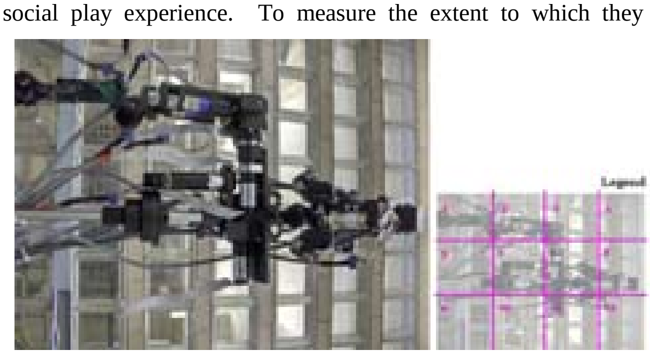

Figure 2. Rotated photograph of Nico used in Survey 1 for identification of body parts.

implicitly viewed Nico as anthropomorphic, this survey posed questions about Nico's attitude, the realism of Nico's behavior, and the quality of the interaction. The three questions dealing with attitude asked whether Nico was attentive ("How well did Nico pay attention to you?"), friendly ("Was Nico friendly?"), and happy (Did Nico seem happy?"). Questions dealing with Nico's behavior included "Were Nico's actions natural?", "Did Nico react appropriately to your actions?", and "Does Nico seem alive?". Lastly, subjects were asked how much they enjoyed interacting with the robot, and given space to write a few sentences describing Nico's positive and negative attributes.

On the same page as the other questions dealing with attitude, subjects were also asked to write a few sentences on the prompt "Please describe Nico's attitude, and **why** you think he acted that way." This open-ended question was designed to solicit each subject's conscious and unconscious beliefs relating to the causes of Nico's behavior. Even if it is entirely clear that a computer program controls Nico, anthropomorphic identification with the robot could lead to higher levels of intention attribution on subject's responses to a broadly defined question. The written prose responses to this question were coded with a method adapted from Malle's F.EX coding scheme [8]. The method depends on examining the reasoning used by respondents to explain the cause of behavior in another agent. It was used in this study to measure the extent to which participants attributed Nico's behavior to an internal intention, which could be considered a projection of human features onto a machine. The original coding scheme was described with a generalized flowchart and a great quantity of examples to clarify the finer points of the system. Those examples were used to develop a new set of examples that would cover the range of possible intention explanations for a robotic agent, and Malle's numeric codes were translated to a four-point scale for the purposes of quantifying the survey

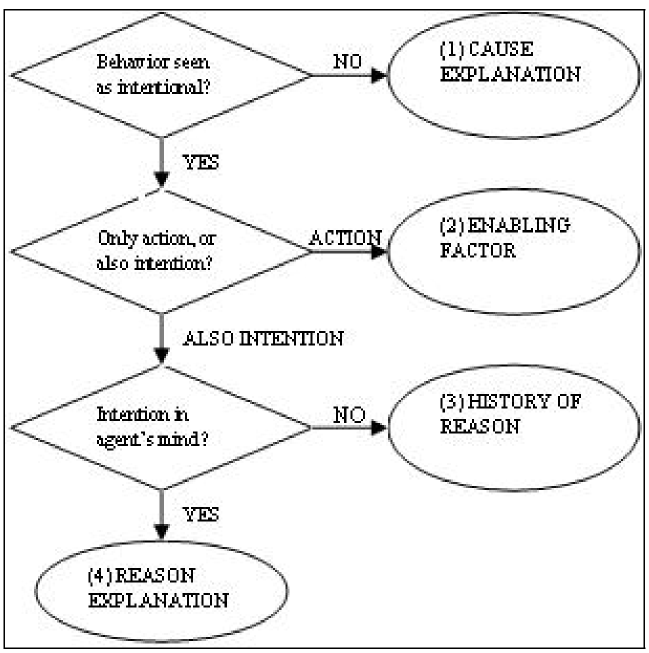

Figure 3. Guide to coding robot intentionality from written descriptions of attitude.

Table 1. Percent of subjects correct on identifying Nico's body parts from a static photograph.

		Control	Experiment
Form Identification		23.1%	70.0%
Rotated Body Identification			
	Arm	46.2%	77.5%
	Hand	45.2%	70.0%
	Neck	23.1%	70.0%
	Mouth	53.8%	55.0%
	Eyes	30.8%	82.5%
	Chest	46.2%	60.0%
Upright Body Identification			
	Arm	76.9%	87.5%
	Hand	92.7%	87.5%
	Neck	84.6%	75.0%
	Mouth	30.8%	42.5%
	Eyes	15.4%	25.0%
	Chest	76.9%	42.5%

responses. An example of a Cause Explanation (1), where an action is completely unintentional, would be "Nico jerked his arm because the shoulder motor froze." An Enabling Factor explanation (2) for Nico would include any appeal to the program as a source of behavior. History of Reason explanations (3) cast Nico as an autonomous agent, but do not cite specific mind content, such as "Nico moved slowly because he was confused." A full-blown Reason Explanation (4) either implies or states explicitly that Nico had a goal in mind, as with "Nico was trying to follow me with her eyes." Figure 3 presents a condensed synopsis of the method, and was used as a guide in coding interpretations of Nico's attitude.

3. RESULTS AND DISCUSSION

The control group consisted of six males and seven females, including 3 science majors (with Psychology and Cognitive Science majors counted as science majors for the purposes of this study), for a total of 13 control subjects. The experimental group had 18 males and 21 females, including 23 science majors, giving a total of 39 subjects that were distributed between the four experiment conditions. Six trials on mode 1, 13 trials on mode 2, 11 trials on mode 3, and 10 trials on mode 4 make up the experimental data analyzed here. Most subjects were in their early twenties, therefore no age analysis was performed. There did not appear to be any strong effects of participants' gender or college major, except that 5% of females in the experimental group viewed Nico as female, while the remainder of the experimental group and the entire control group was split evenly between viewing Nico as male, or as having no gender. This slight effect of the interaction experience might be an indicator that playing with Nico in person was necessary for subjects to project their personal qualities onto an animated robot.

Table 2. Mean ratings on 8 measures of social experience.

Condition	No Tracking	Smooth Tracking	Unsmooth Tracking	Avoidance
Attention	1.83	2.69	3.27	2.50
Friendly	1.80	2.85	3.20	3.25
Happy	1.33	2.73	2.25	2.57
Intention	1.33	2.08	2.64	2.40
Natural	2.00	3.00	2.64	2.90
Appropriate	1.83	2.31	3.40	2.50
Alive	1.00	2.77	2.36	1.90
Enjoyment	1.80	2.75	3.18	3.20

3.1 Anthropomorphic Identification

In the panoramic photo, 77% of experimental subjects identified Nico as an anthropomorphic form, while only 30% of control subjects succeeded on this measure. ANOVA showed this result to be significant, $F(1, 51) = 10.29$, $p < .003$. The 23% of experimental subjects who did not list Nico in their answer may or may not have omitted the answer as a conscious choice. It is unclear whether the explanation of the anthropomorphism concept was not detailed enough, or whether some subjects who had played with Nico simply decided that the robot had no anthropomorphic form. Nico's lack of a right arm and lower body was cited in a number of survey responses as detracting from structural form.

3.2 Body Identification

In the identification using a rotated picture, control subjects failed on average about 59% of the time, while experimental subjects failed around 42% of the time (see Table 1). With the upright picture, the control group improved to a slightly better success rate than the experimental group, failing at 37% as opposed to an average of 40%. It is unclear why the control group had more success than the experimental group once the picture was oriented logically. One possibility is that Nico's structural anthropomorphic form was not memorable enough for experimental subjects, perhaps due to the very undifferentiated mechanical character of Nico's body. Much of the failure in both groups was not due to the rotation of the photograph, but rather to the perception of Nico's eyes being higher on the head and the belief that he had a mouth. Despite the fact that many experimental subjects noted and were frustrated by Nico's lack of sound production, many still answered that Nico had a mouth, possibly indicating attribution of a human form even when the form was absent in the actual physical structure.

The identification of body parts was used as a measure of recognition since a highly anthropomorphic form should have parts that are widely seen as corresponding with human appendages. The higher success of experimental subjects could be attributed to their knowing which way the rotated photograph should be interpreted, their having recognized the individual parts during the interaction, or both.

3.3 Social Group Indicators

Regardless of experimental condition, nearly all subjects were split evenly on the question of Nico's gender, with half choosing "Male," and half adamantly stating their inability to choose a gender when no gender cues were provided. Many subjects noted that the name "Nico" sounds like a male diminutive, and that Nico's angularity and dark coloring contribute to an air of masculinity. Thus, it is clear that Nico's behavior was neither strongly gendered nor characterized enough for female subjects, who made up just over half the subject population, to strongly project their own gender and personal characteristics onto the robot.

In the control group, the 54% majority believed Nico to represent a child. In the experimental group, the distribution of age responses drifted downward slightly, with the 45% majority choosing toddler. Experimental subjects cited reasons such as Nico's relative immobility, pre-verbal behavior, and confused attitude. It is likely that the simple nature of Nico's behavior during the interaction led experimental subjects to perceive Nico as a younger, less developed being.

3.4 Measures of Social Experience

The eight measures for the interaction experience were intended to fall roughly into three categories: positive affect (e.g. attention, friendliness, and happiness) contingency (e.g. naturalness, appropriate reactivity, and seeming alive), and enjoyableness. The additional intention measure was derived from coding written opinions of Nico's attitude, and has the same range of values as the other ratings (see Table 2).

As expected, the no tracking condition received low ratings across the board, since a static head should detract from the observer's engagement relative to any kind of moving head. The smooth tracking condition showed a marked improvement over no tracking, but still rated mostly below average. Unsmooth tracking received the most above-average ratings, for attention, friendliness, appropriateness, and enjoyment, but the avoidance condition surpassed unsmooth tracking on friendliness and enjoyment. This result implies that less ordinary head movements were more pleasing to subjects than the head behavior that most closely mimicked normal head movement.

Surprisingly, measures tended to correlate more highly between categories rather than within categories. Thus, within the positive affect category, happiness and attention were correlated 0.86 and 0.83 respectively to friendliness, but friendliness, intention, and enjoyment, three measures from different categories, were correlated to each other with r-values between 0.97 and 1.00 (see Figure 4). These high correlations suggest that the categories used in our survey do not accurately reflect how subjects' attitudes are actually grouped during the social interaction. That is, the same features of a humanoid robot's behavior trigger perceptions of intentionality, enjoyment, and friendliness on the part of the observer, and these should form one category.

The most significant result, measured by ANOVA, was on the attention measure, $F(2, 36) = 4.20$, $p < .02$. The measures that correlated highly with intentionality were only moderately significant: on enjoyment, $F(3, 34) = 2.38$, $p < .09$; on friendliness, $F(3, 32) = 1.86$, $p < .16$; and on intention, $F(3, 36) = 2.20$, $p < .11$. Naturalness and happiness were also correlated near 1.00, but turned out to be insignificant, with $p > .2$. The fact that these weak effects were correlated suggests that subjects may have had no strong feeling on the measures, but that they still related the answers to both questions very closely.

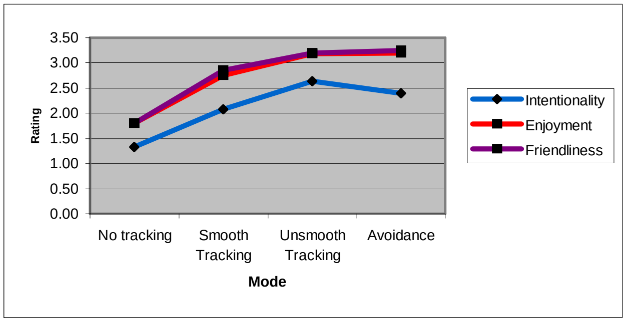

Figure 4. Intentionality measure correlations.

4. CONCLUSIONS

It is clear that variation in these mannerisms of head movement play with the basic mechanisms that subjects used to evaluate the behavior of an obviously non-living machine. Even our implementation of only a very small set of contingent behaviors was perceived much more positively when the robot's head had a more active, if less realistic, characterization and implementation.

Our results were consistent with those of Vinayagamoorthy, Steed, and Slater's conclusion that a virtual character should have comparable levels of visual and behavioral realism to enhance user responses [10]. Since Nico's appearance was not very life-like, experimental conditions with more realistic movements resulted in poorer user responses.

None of the social experience ratings for any condition consistently scored above average, perhaps because subjects started with unrealistic expectations of robots from science fiction. But the tracking mode where Nico actively turned away from the subjects proved to be the most enjoyable, because it established an implicit game of peek-a-boo with the subject.

Thus, in making humanoid robots useful to naïve users, realism should perhaps not be the ultimate goal. This experiment has demonstrated that users may prefer exaggerated, caricatured behavior in a robot over realistic human behavior. Even though caricatures are obviously less realistic and unnatural, they are more immediately lovable. Although subjects accurately perceived Nico with smooth tracking as the most natural, the experimental group preferred playing with Nico with unsmooth tracking and avoidance, where the robot's actions were characterized with stronger indications of intention. The overriding impression of Nico's more active head movements was that the robot was busily trying to compute a judgment, while less head movement made Nico seem passive. Across all conditions, the scripted arm gestures aroused initial interest in the naïve visitor.

While the survey method we used to examine the subjects' post-hoc impressions gave moderate effects, a more direct measure of the subjects' experience would be more desirable. One possibility is videotaping the interaction, and coding the recordings for eye contact and other types of body language. Additional work that uses video coding in conjunction with the survey method could also confirm the validity of the coding scheme used here to gauge the perception of intentionality,

5. ACKNOWLEDGMENTS

We thank Fred Shic for his guidance on developing and debugging code for Nico and our experimental subjects for their participation.

6. REFERENCES

[1] Bartneck, C., and Forlizzi, J. 2004. Shaping human-robot interaction: Understanding the social aspects of intelligent robotic products (pp. 1731 – 1732). CHI04 Extended Abstracts, Vienna, Austria, April.

[2] Breazeal C., and Scassellati B. 1999. How to Build Robots that Make Friends and Influence People. Presented at the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-99), Kyongju, Korea.

[3] Caporael, L.R., and Heyes, C.M. 1997. Why anthropomorphize? Folk psychology and other stories. In Mitchell, R.W, Thompson, N.S, and Miles, H.L. eds., *Anthropomorphism, Anecdotes, and Animals*. Albany, New York: State University of New York Press.

[4] DiSalvo, C., and Gemperle, F. 2003. From Seduction to Fulfillment: The Uses of Anthropomorphic Form in Design. *Designing Pleasurable Products and Interfaces*. Pittsburgh, Penn.: ACM Press.

[5] DiSalvo, C., Gemperle, F., and Forlizzil, J. Imitating the Human Form: Four Kinds of Anthropomorphic Form. Unpublished manuscript. Accessed April 2005 from <http://www.anthropomorphism.org/pdf/Imitating.pdf>.

[6] Kiesler, S. 2005. <http://www.anthropomorphism.org>. Carnegie Mellon University.

[7] Kiesler, S., and Goetz, J. 2002. Machine trait scales for evaluating mechanistic mental models. Unpublished manuscript. Accessed from <http://www.peopleandrobots.org>.

[8] Malle, B.F. 2000. F.EX: A Coding Scheme for Folk Explanations of Behavior, Vers 4.0. University of Oregon. Accessed April 2005 from <http://hebb.uoregon.edu/04-02tech.pdf>.

[9] Mori, M. 1970. Bukimi No Tani [the uncanny valley]. *Energy*, 7:33-35.

[10] Vinayagamoorthy, V., Steed, A., Slater, M. 2005. Building Characters: Lessons Drawn from Virtual Environments. In *Towards Social Mechanisms of Android Science*, Stresa, Italy, pp. 119-126.