Playing Catch with Robots: Incorporating Social Gestures into Physical Interactions

Elizabeth J. Carter[1], Michael N. Mistry[2], G. Peter K. Carr[3], Brooke A. Kelly[3] and Jessica K. Hodgins[1,3]

***Abstract*— For compelling human-robot interaction, social gestures are widely believed to be important. This paper investigates the effects of adding gestures to a physical game between a human and a humanoid robot. Human participants repeatedly threw a ball to the robot, which attempted to catch it. If the catch was successful, the robot threw the ball back to the human. For half of the cases in which the catch was unsuccessful, the robot made a physical gesture, such as shrugging its shoulders, shaking its head, or throwing up its hands. In the other half of cases, no gestures were produced. We used questionnaires and smile detection to compare participants' feelings about the robot when it made gestures after failure versus when it did not. Participants smiled more and rated the robot as more engaging, responsive, and humanlike when it gestured. We conclude that social gesturing of a robot enhances physical interactions between humans and robots.**

I. INTRODUCTION

As robots become more integrated into daily life, it is increasingly imperative for them to be able to interact with humans socially as well as physically. Specifically, robots that provide entertainment and household support will need to go beyond physically performing tasks (e.g., doing laundry, baking cookies) and provide social feedback to the users to increase the convenience of their interactions and ensure the robots' use [9]. In this paper, we describe research that investigates how the addition of social gestures to a physically based task affects people's opinions of a humanoid robot used for entertainment.

We elected to have the robot play a game of catch with humans because it is a multi-step, interactive behavior that requires observable realtime processing and builds upon the robot's humanoid characteristics. Additionally, it is a form of interaction that provides a physical connection with the robot yet keeps participants at a safe distance. In our paradigm, the human participant threw a ball to the robot, which was able move its left hand to catch the ball if the ball intersected a particular volume of space near the arm. Upon catching the ball, the robot would toss it back in the direction of the human. Because of inconsistencies in the humans' throws, the robot had to constantly adapt its hand location in an effort to catch the ball. Even the most skilled humans could not consistently throw within the robot's working range on every trial. Thus, it was clear to the users that the robot was assessing the ball position and making decisions in realtime in direct response to the environment.

Previous research has examined catching and throwing in human-robot pairs [30], [20], but it has focused on attaining ideal physical interactions without significant examination of how the addition of social behaviors, such as gestures, might change the humans' subjective experience of the interaction. Therefore, we created head motions that followed the trajectory of the ball in the air and various large-scale movements that occurred contingently in the case of catching failure, including shoulder shrugging, head shaking, and throwing up hands in despair. We assessed how the addition of these social actions affected the overall interaction experience in two ways. We surveyed participants about the interactive qualities of the robot, such as engagement, responsiveness, and friendliness. We also examined participants' facial expressions during the interaction for smiles and laughs.

II. RELATED WORK

Robots have been playing catch with humans for years: it is a dynamic, physical interaction that does not require physical contact with the human while also a technologically complex task. The robots have relied upon computer vision to predict trajectories [17] or visual servoing techniques [26], [10]. Catching can avoid [3], [30] or involve grasping movements [32]. We recently described our own catching robot system for human-to-robot partner juggling [20], which we have modified for the present work. We sought to make people's interactions with the robot more positive through the addition of head and arm gestures to the robot's behavior.

Bodily gestures can be used for communication in place of or in conjunction with speech (for a review, see [13]). Emblematic gestures convey information without the need for concurrent speaking; cultures have agreed-upon meanings for these specific motions [18]. Common emblems include head nods, hand waves, and shoulder shrugs. Conversational gestures are synchronized with speech and can relate to its meaning, but do not occur in its absence. These can include beat gestures as well as more complex movements. Finally, adapters include manipulation of the self, another, or an object without communicative intent during speech.

Previous research has investigated how physical movements, including gestures, change viewer perception of synthetic agents. Humans tend to respond socially to non-human objects like computers [28], and this can be generalized to robots. Moreover, the level of social response can be

This work was supported by Disney Research.

[1]Robotics Institute, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213 USA lizcarter@cmu.edu

[2]School of Computer Science, University of Birmingham, Edgbaston, Birmingham, B152TT UK m.n.mistry@bham.ac.uk

[3]Disney Research Pittsburgh, 4720 Forbes Avenue, Pittsburgh, PA 15213 USA peter.carr, brooke.kelly, jkh@disneyresearch.com

manipulated by agent characteristics. One goal of human-robot and human-computer interaction is to determine how to design robots and agents that respond to humans in a social way, assuming it is appropriate given the context [9], [12]. Prior research suggests that gestures can provide social cues that assist in improving humans' perception of robots.

Some robots have relied upon head motion to support interactions. Sidner and colleagues [33] found that using head movement to track the participant during a demonstrative conversation with a penguin robot elicited more attention to the robot, and participants found the movements appropriate. In another study [21], a humanoid robot took turns drumming with a participant with or without head movements. Participants believed that they performed better in the absence of the head motion; however, they enjoyed the interaction that had head motion more. Other research [14] examined the use of head nods, eye blinks, and gaze shifts in various combinations by a robotic head and determined that robots that exhibited head nods were perceived as more engaged and likable.

Additional research has examined the use of head and arm movements in combination to convey gestural information. Prior research has demonstrated that head and arm movements of a humanoid robot can affect user impressions of the personality dimensions of that robot, including introversion/extraversion and thinking/feeling [19]. Moreover, the motion affected ratings of other characteristics, such as excitement, pleasantness, likability, and interestingness. In another study, adding head and arm movements to a small telepresence robot increased users' feelings of engagement, enjoyment, and cooperation with the robot [1]. In a video simulation study, Takayama and colleagues [34] found that an animated robot that changed its body and arm postures to acknowledge failure at a task were viewed as more smart and confident than those that did not physically acknowledge the failure. More recently, Salem and colleagues [31] had participants follow instructions given by a small, humanoid robot to unpack a box and put away its contents. The robot gave verbal instructions either without referential gestures, with correct gestures, or with partly incorrect gestures. Gestures, whether correct or incorrect, increased ratings of humanlikeness, shared reality, and likability. The highest ratings were assigned to the robot when some gestures were incorrect, suggesting that errors can make robots seem more human.

These studies suggest that adding nonverbal behaviors, including gestures, to a robot can improve the users' interaction experiences with that robot. Our study explores how the incorporation of gestures during a game of catch affects the users' ratings of the robot. Our study expands on current research by manipulating gestures in a more natural, physically interactive task and having the participants evaluate the robot's traits and interaction as a whole, as opposed to explicitly having them evaluate the gestures.

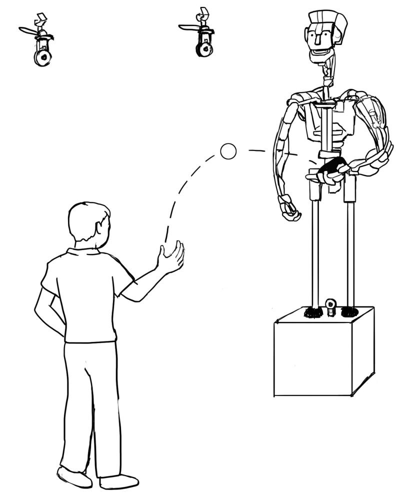


Fig. 1. The human participant stands approximately 1.5m in front of the humanoid robot. The human and robot throw a ball back and forth. The ball is tracked using two ceiling-mounted cameras. A camera mounted on the base of the robot monitors participants' faces.

A. Hypotheses

We created two experimental conditions for our research: the robot either did or did not give gestural feedback after catching failures. Our first research question (Q1) was whether the introduction of social gestures into the robot's behavior would improve participants' ratings of the characteristics of the robot. We collected ratings of how humanlike, engaging, competent, responsive, friendly, and attractive the robot was after games with and without gestures. We hypothesized (H1) that gestures would improve these ratings, particularly for humanlikeness, engagement, responsiveness, and friendliness. For our second question (Q2), we examined users' facial expressions while they played catch with the robot to determine whether the robot's social displays could elicit increased social displays from the users. We expected (H2) that the participants would show more social displays (in this case, smiles) when the robot was also making social displays.

III. APPARATUS

Our experimental setup includes a human participant, a humanoid robot, and a computer vision system (Fig 1).

A. Robot Hardware

Our catching robot is a Walt Disney Imagineering A100 Audio-Animatronics figure (Fig. 4). This platform is commonly employed in theme parks and other entertainment venues. The robot is 1.8m tall and stands with feet fixed on a 0.6m base. The robot contains a total of 38 hydraulically actuated degrees-of-freedom (DOF). For throwing and catching, we use the left arm (seven DOF plus one DOF for each of the five fingers) and the torso (two DOF for bending forward and twisting). For social gestures (including head turning to follow the ball), we additionally use the right arm, shoulder, neck and eyes. We augmented the left hand of the robot with a plate to cover and protect the finger actuators

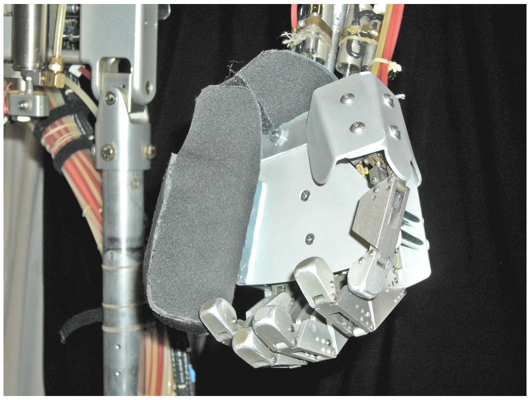


Fig. 2. Detail of the robot's catching hand. Black foam has been placed on the back and side of the palm to form a basket shape for catching. The palm is approximately 10cm square.

Fig. 3. Each camera performs ball detection (red circle) independently. The 3D coordinate of the ball (blue curve) is then triangulated from the pair of 2D image locations. When a throw is detected, a Kalman filter (green curve) processes the raw 3D positions and predicts a 3D catching location (pink square) and time to impact.

and a foam rim to provide a more cup-like shape suitable for catching (Fig. 2). The fingers of the hand do not have the range of motion or capability for grasping the ball and instead are used to form a basket shape while catching and to extend during throwing. Further information about the robot and its control system is described in previous work [20].

B. Computer Vision

Previously, we used a Asus Xtion PRO LIVE color/depth camera for ball tracking [20]. However, we found the latency too high for the present work. Thus, we replaced the Asus camera with two GigE machine vision cameras for tracking the ball in 3D space. The entire process (image processing, network communication and triangulation) took approximately 33ms to complete.

A Kalman filter was used to predict the catching location for the robot and anticipated time to impact (Fig 3). The filter was reinitialized whenever the system detected the ball moving with a sufficient upwards and forwards velocity and originating reasonably close to the participant's location. With each new observed 3D location, the filter would update its anticipated catching location and time to impact, and examine the tracking residual. If the tracking residual was significant after the time of impact, our residual system would determine whether the ball was caught or had bounced off the hand. In our previous vision system [20], due to latency, we were only able to determine a missed ball by checking if the ball position fell below the catching plane. However, by examining the tracking residual, we were able to obtain an instantaneous estimate of whether or not a ball was "almost" caught (i.e., hitting the hand but bouncing off, as opposed to missing the hand completely). Thus we were able to cue an appropriate gesture that acknowledged a near catch versus complete miss.

A third GigE machine vision camera was used to monitor the facial expressions of the participant. Images of the participant were captured at 30 frames per second and recorded to disk. The video recordings of the participants' facial expressions were manually annotated to identify the release point of each throw. The participants' faces were automatically detected and analyzed for smiles over the next 5 seconds (150 frames) of video using the Fraunhofer face detector and analysis engine [22].

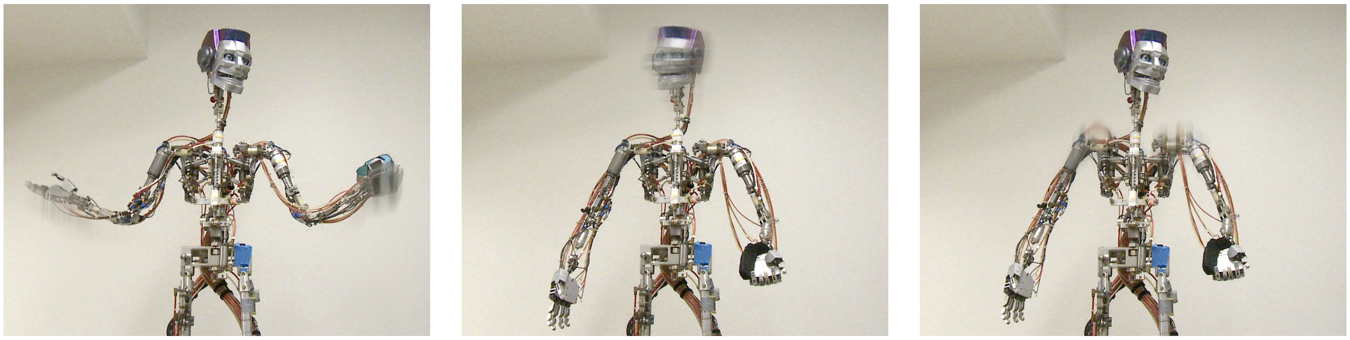
IV. Method

For the 'Gestures' condition, we integrated multiple motion features into the robot's behavior. First, the robot followed the ball with his head, watching the ball as it was tossed back and forth. Three gestures were created for when the robot missed catching the ball. In the case of a near miss, the robot would either throw its hands up in the air for the "Raise Arms" gesture or move his shoulders up and down for the "Shrug Shoulders" gesture. If the throw went far out of the catching range of the robot, it would shake its head back and forth, the "Shake Head" gesture. We selected these gestures because they are emblematic and can be understood in the absence of speech, i.e., throwing up one's hands as a symbol of despair, shrugging when one doesn't know what to do, and shaking one's head to say no. Moreover, gestures that acknowledge a robot's failure have been demonstrated previously to affect user experience in a non-interactive task [34], and this subset of failure gestures could be performed by our robot in a comprehensible, efficient manner.

A. Human Interaction Study

1) Participants: Thirty adults (18 male, 12 female) aged 19 to 64 years (average 27.83, standard deviation 10.75) participated in this research. Participant recruitment was performed via an online participation pool open to the local community. We required that the participants have normal or corrected-to-normal vision and be able to throw and catch. The research was approved by the Institutional Review Board and participants were compensated for their time.

2) Procedure: This experiment was performed in a 'within-subjects' design so that each participant played with the robot both with and without gestures. The order of the conditions (henceforth referred to as 'Gestures' and 'No Gestures') was randomized such that half of the participants had 'Gestures' first and the other half had 'No Gestures' first. Two experimenters were present at all times to assist



Raise Arms | Shake Head | Shrug Shoulders

Fig. 4. If the robot was unable to catch the ball, it optionally played one of three gestures: raising arms, shaking head or shrugging shoulders.

with the study, one of whom was responsible for ensuring that the robot was performing properly. At the onset of the study, the second experimenter explained to each participant that the robot was a left-handed catcher and responded best to underhand tosses. Then, the participant threw a minimum of twenty times to the robot. (In some cases, the robot miscounted and took an additional turn.) In either condition, if the robot succeeded in catching, the robot would toss the ball back. In the 'No Gestures' condition, if the robot failed to catch the ball, the robot simply returned to its starting ready-to-catch pose, and the experimenter returned the ball to the participant. In the 'Gestures' condition, if the toss was a near miss (the ball deviated from its parabolic trajectory because it hit the robot's hand), the robot performed either the 'Raise Arms' or 'Shrug Shoulders' gesture with equal probability (Fig 4). If the toss was a complete miss, the robot performed the 'Shake Head' gesture. After the gesture was complete, the robot returned to its ready-to-catch pose, and the experimenter returned the ball to the participant. After each set of throws, the participant filled out a brief survey (a modified, shortened adaptation of the GODSPEED questionnaire [2]) and the gesture condition was switched. The participant then threw the ball to the robot for another set and filled out the same survey again. Finally, the participant answered a few additional questions about his or her background and experience to ensure that he or she did not need to be excluded from analyses for high familiarity with robots. No participants were excluded.

3) Measures: After each block of throws, participants responded to the following questions on a 5-point scale:

This time the robot was...

- very humanlike (1) to very unhumanlike (5)
- very engaging (1) to very unengaging (5)
- very competent (1) to very incompetent (5)
- very responsive (1) to very unresponsive (5)
- very friendly (1) to very unfriendly (5)
- very attractive (1) to very unattractive (5)

B. Analysis

We performed a repeated-measures, multivariate analysis of variance (MANOVA) with one independent variable (condition: 'Gestures' and 'No Gestures') and six dependent variables (humanlike, engaging, competent, responsive, friendly, and attractive ratings) in order to assess the effect of gestures on participant opinions of the robot. One participant was not included in this analysis because she did not complete all questions of the survey.

Faces were detected with 91.7% recall, and when a face was detected, the smile score ranged from 0 to 100. The smile data exhibited bimodality and was not normally distributed: the majority of the time, participants had a neutral, non-smiling expression, but occasionally the smile score would be significant.

We formulated the null hypothesis that the smiles scores in 'Gestures' and 'No Gestures' come from the same unknown distribution. Because the global smile scores were not normally distributed, we used the Wilcoxon signed-rank test to determine whether the smile scores during 'almost' and 'unsuccessful' catches with and without gestures came from the same distribution (recall that no gesture was played for a successful catch).

V. Results

In all, there were 1199 throwing trials. As mentioned previously, the robot's catching performance was automatically classified by the vision system's Kalman filter tracking error into three categories: (1) a successful catch, (2) an almost successful catch (where the ball deviated from its parabolic trajectory because it hit the robot's hand), and (3) an unsuccessful catch (because the ball did not deviate from its expected trajectory). The breakdown of the 1199 throws is listed in Table I. There was no significant difference in the number of successful throws between the 'Gestures' and 'No Gestures' condition, $t = 0.41$, $p = 0.68$. Performance ranged from 0 to 16 successful throws per condition.

A. Questionnaires

Mean ratings and standard deviations for each of the six questions by condition are presented in Table II. For all questions combined, there was not a significant effect of condition, $F(6,23) = 1.77$, $p = 0.15$. However, some individual questions had significantly different results for the 'Gestures' and 'No Gestures' conditions. With gestures, the robot was rated as significantly more humanlike than without gestures ($F = 6.947$, $p = 0.01$, observed power = 0.72). Gestures also significantly improved ratings of responsiveness ($F = 3.88$,

TABLE I
THROWING OUTCOMES

Condition	Count	Per Participant Mean	St. Dev.
Successful	525	17.5	6.3
No Gestures	267	8.9	3.4
Gestures	258	8.6	4.0
Almost	169	5.6	3.0
No Gestures	77	2.6	1.7
Gestures	92	3.1	2.0
Unsuccessful	505	16.8	6.9
No Gestures	256	8.5	4.1
Gestures	249	8.3	4.0

TABLE II
MEANS AND STANDARD DEVIATIONS OF QUESTIONNAIRE RESPONSES BY CONDITION. ASTERISKS INDICATE SIGNIFICANCE AT P < 0.05.

Question	No Gestures		Gestures	
	Mean	St. Dev.	Mean	St. Dev.
Humanlike*	2.55	0.74	3.03	0.82
Engaging*	2.59	1.09	3.07	1.03
Competent	2.93	0.96	3.14	0.88
Responsive*	2.34	0.90	2.86	0.88
Friendly	2.41	1.15	2.72	0.96
Attractive	2.59	1.02	2.76	0.99

p = 0.01, power = 0.75) and how engaging the robot was (F = 3.38, p = 0.05, power = 0.51). A trend was identified such that ratings of friendliness were improved by gestures (F = 2.77, p = 0.11, power = 0.36); however, further research is needed to examine this dimension. There were no significant differences across conditions for ratings of how competent or attractive the robot was (F = 1.00, p = 0.33 and F = 1.33, p = 0.26). We believe that these results arise from the consistency across conditions in the robot's ability to catch successfully and his appearance.

B. Facial Expressions

The histograms of the smile scores with and without gestures are shown in Figure 5. Enabling gestures resulted in a noticeable decrease of weak smile scores (≈ 5) and increase in strong smile scores (≈ 95). The histograms in Figure 5 are biased towards participants with fewer successful catches—i.e., participants with more 'almost' and 'unsuccessful' catches contributed more data points for smile scores with and without gestures. However, equivalent histograms that normalize the number of samples from each participant show the same trend, but with slightly larger magnitude: the decrease in weak smiles is more pronounced, as is the increase in strong smiles. Due to the high performance of a small subset of participants at throwing, we eliminated 5 participants who had 12 or more successful throws per condition from the statistical analysis. These participants did not have the opportunity to see the full range of gestures. Smile scores were created for each 'almost' and 'unsuccessful' catching trial by averaging the scores across the subsequent 150 frames of video, as this gave participants an adequate period of time to view the gesture (if displayed)

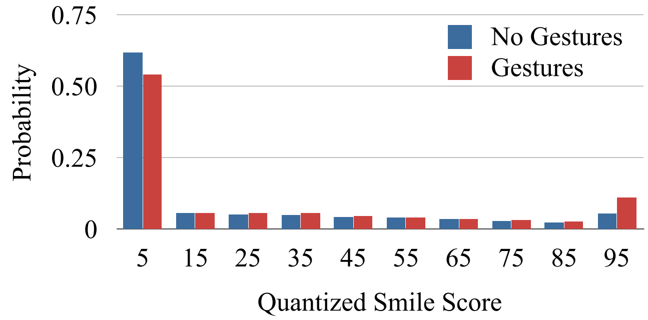


Fig. 5. The histograms of smile scores for gestures disabled and enabled during 'almost' and 'unsuccessful' catches exhibit similar distributions. However, when gestures are enabled, there is a noticeable and significant decrease in weak smiles and an increase in strong smile scores.

and react before the next trial. We performed a Wilcoxon signed-rank test (p=0.02) and found that the smile scores in the 'Gestures' condition are significantly different (at a 95 percent confidence level) compared to smiles in the 'No Gestures' condition.

VI. DISCUSSION

In cooperative tasks involving humans, gestures are a natural way for robots to exhibit their understanding of the world. For instance, in our scenario, the A100 robot turns its head to follow the ball (demonstrating awareness of the ball's location in the world). Similarly, when the robot was unable to catch the ball in the Gestures condition, it acknowledged its failure through well-understood human body gestures. Our study demonstrated how the incorporation of these gestures made the throwing and catching interaction more positive for the human participants. The survey results indicated that gestures made the robot seem more humanlike, engaging and responsive, presumably because it demonstrated awareness of failure. The gestures also elicited increased smile scores from the participants, implying that the gestures successfully conveyed an appropriate and somewhat humorous acknowledgment of failure by the robot.

Our current study showed how gestures made the throwing and catching interaction more enjoyable for the human participants. In future work, we plan to explore whether gestures can be used to improve the performance of a cooperative task (such as maximizing the number of consecutive successful throws and catches between the human and robot). If a robot fails to perform a particular task, but is cognizant of why the failure occurred, the robot can acknowledge this through a gesture. Because our vision system can also determine if the toss was a complete miss, or if the ball was nearly caught before bouncing out of the hand, the robot can perform a gesture that directs blame appropriately.

Different gestures indicate different levels of awareness. For example, one participant commented that the robot should look at him more often instead of only looking at the ball during the 'Gestures' condition. Having the robot look at the participant while waiting for the ball to be thrown and then follow the ball while it is traveling through the air demonstrates a more complete understanding of the task. With reliable realtime face tracking, we should be able to modify the robot's behavior to include attention to the participant's face and apparent eye contact. It would be interesting to see whether these changes in head gesture make the robot seem more competent.

In the current study, the robot only made gestures after failing to catch the ball. However, we would also like to explore whether excited and/or happy gestures after a successful catch have similar or even more profound effects on the human participants' opinions of the interaction if we could prevent the ball from dropping out of the hand. If smile detection could be performed in realtime, the robot could adapt its repertoire of gestures to suit each participant. Additionally, future work should address how the use of random or inappropriate gestures would affect people's experience of their interactions with the robot. Such research would clarify whether movement alone can make robots seem more humanlike, engaging, and responsive during this type of exchange.

Finally, other modes could be added, such as speech responses. With these additional capabilities, the robot could be equipped with behaviors to help coach participants in throwing the ball. These feedback cues should result in ball trajectories that pass through the robot's working volume more often and result in an improved catching success rate.

ACKNOWLEDGMENT

We thank Jens Kober for much of the original work on the A100 catching system. We are grateful to Matthew Glisson and Rachel Browne for their assistance in performing this research. We also thank our participants. This work was funded by Disney Research.

REFERENCES

[1] S.O. Adalgeirsson and C. Breazeal MeBot: A robotic platform for socially embodied telepresence. In *Proc. Int. Conf. on Human-Robot Interaction*, 2010.

[2] C. Bartneck, E. Croft, D. Kulic, and S. Zoghbi. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. In *Int. J. Social Robotics*, 2009.

[3] G. Bätz, A. Yaqub, H. Wu, K. Kühnlenz, D. Wollherr, and M. Buss. Dynamic manipulation: Nonprehensile ball catching. In *Proc. Mediterranean Conf. on Control and Automation*, 2010.

[4] B. Bäuml, O. Birbach, T. Wimböck, U. Frese, A. Dietrich, and G. Hirzinger. Catching flying balls with a mobile humanoid: System overview and design considerations. In *Proc. Int. Conf. on Humanoid Robots*, 2011.

[5] B. Bäuml, T. Wimböck, and G. Hirzinger. Kinematically optimal catching a flying ball with a hand-arm-system. In *Proc. Int. Conf. on Intelligent Robots and Systems*, 2010.

[6] O. Birbach and U. Frese. Estimation and prediction of multiple flying balls using probability hypothesis density filtering. In *Proc. Int. Conf. on Intelligent Robots and Systems*, 2011.

[7] O. Birbach, U. Frese, and B. Bäuml. Realtime perception for catching a flying ball with a mobile humanoid. In *Proc. Int. Conf. on Robotics and Automation*, 2011.

[8] G. Collier and D. DiCarlo. *Emotional expression.* Lawrence Erlbaum Associates, Inc, 1985.

[9] K. Dautenhahn. Socially intelligent robots: Dimensions of human–robot interaction. *Phil. Trans. Royal Society B: Biol. Sci.*, 362(1480):679–704, 2007.

[10] K. Deguchi, H. Sakurai, and S. Ushida. A goal oriented just-in-time visual servoing for ball catching robot arm. In *Proc. Int. Conf. on Intelligent Robots and Systems*, 2008.

[11] U. Frese, B. Bäuml, S. Haidacher, G. Schreiber, I. Schaefer, M. Hähnle, and G. Hirzinger. Off-the-shelf vision for a robotic ball catcher. In *Proc. Int. Conf. on Intelligent Robots and Systems*, 2001.

[12] J. Goetz, S. Kiesler, and A. Powers. Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *Proc. Int. Workshop on Robot and Human Interactive Communication*, 2003.

[13] S. Goldin-Meadow. The role of gesture in communication and thinking. *Trends in Cognitive Sciences*, 3(11):419–429.

[14] J. Hall, T. Tritton, A. Rowe, A. Pipe, C. Melhuish, and U. Leonards Perception of own and robot engagement in human-robot interactions and their dependence on robotics knowledge *Robotics and Autonomous Systems*, in press.

[15] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision.* Cambridge University Press, second edition, 2004.

[16] T. P. Hettmansperger and J. W. McKean. *Robust nonparametric statistical methods*, Volume 119 of *Monographs on statistics and applied probability*. CRC Press, Boca Raton, FL, 2nd Edition, 2011.

[17] B. Hove and J.-J. E. Slotine. Experiments in robotic catching. In *Proc. American Control Conf.*, 1991.

[18] A. Kendon. Some recent work from Italy on quotable gestures ('emblems'). *J. Linguist. Anthropol.* 21:72–93.

[19] H. Kim, S. S. Kwak, and M. Kim. Personality design of sociable robots by control of gesture design factors. In *Int. Symp. on Robot and Human Interactive Communication*, 2008.

[20] J. Kober, M. Glisson, and M. Mistry. Playing catch and juggling with a humanoid robot. In *Proc. Int. Conf. on Humanoid Robots*, 2012.

[21] H. Kose-Bagci, K. Dautenhahn, D. S. Syrdal, and C. L. Nehaniv. Drum-mate: interaction dynamics and gestures in human–humanoid drumming experiments. *Connection Science*, 22(2):103–134, 2010.

[22] C. Küblbeck, T. Ruf, and A. Ernst. A modular framework to detect and analyze faces for audience measurement systems. In *GI Jahrestagung*, pages 3941–3953, 2009.

[23] R. Lampariello, D. Nguyen-Tuong, C. Castellini, G. Hirzinger, and J. Peters. Trajectory planning for optimal robot catching in real-time. In *Proc. Int. Conf. on Robotics and Automation*, 2011.

[24] R. Lippa. The nonverbal display and judgment of extraversion, masculinity, femininity, and gender diagnosticity: A lens model analysis. *Journal of Research in Personality*, 32(1):80–107, 1998.

[25] A. Mehrabian. *Silent messages.* Wadsworth, 1971.

[26] R. Mori, K. Hashimoto, and F. Miyazaki. Tracking and catching of 3D flying target based on GAG strategy. In *Proc. Int. Conf. on Robotics and Automation*, 2004.

[27] L. Moshkina and R. C. Arkin. Human perspective on affective robotic behavior: A longitudinal study. In *Proc. Int. Conf. on Intelligent Robots and Systems (IROS)*, 2005.

[28] B. Reeves and C. Nass. *The Media equation: how people treat computers, television, and new media.* Cambridge University Press, 1997.

[29] R. E. Riggio, R. Lippa, and C. Salinas. The display of personality in expressive movement. *Journal of Research in Personality*, 24(1):16–31, 1990.

[30] M. Riley and C. G. Atkeson. Robot catching: Towards engaging human-humanoid interaction. *Auton. Robots*, 12(1):119–128, 2002.

[31] M. Salem, F. Eyssel, K. Rohlfing, S. Kopp, and F. Joublin. To err is human(-like): Effects of robot gesture on perceived anthropomorphism and likability. *Int. J. of Social Robotics*, 5(3):313–323, 2013.

[32] T. Senoo, Y. Yamakawa, S. Mizusawa, A. Namiki, M. Ishikawa, and M. Shimojo. Skillful manipulation based on high-speed sensory-motor fusion. In *Proc. Int. Conf. on Robotics and Automation (ICRA)*, 2009.

[33] C. Sidner, C. Kidd, C. Lee, and N. Lesh Where to look: A study of human-robot engagement In *Proc. Int. Conf. on Intelligent User Interfaces (IUI)*, 2004.

[34] L. Takayama, D. Dooley, and W. Ju Expressing thought: Improving robot readability with animation principles. *Proc. of the 6th Int. Conf. on Human-Robot Interaction*, 2011.