Robot Humor: How Self-Irony and Schadenfreude Influence People's Rating of Robot Likability

Nicole Mirnig, Susanne Stadler, Gerald Stollnberger, Manuel Giuliani, Manfred Tscheligi

Abstract— Humor in robotics is a promising, though not yet significantly researched topic. We performed a user study exploring two different kinds of laughter. In our study, participants observed a robot-robot interaction where an iCat and a NAO robot exhibited different laughing behavior. While NAO laughed at itself (self-irony), the iCat laughed at NAO (Schadenfreude[1]). Our participants watched four turns of the same robot-robot interaction, with either NAO or the iCat laughing, both robots laughing, or no robot laughing (baseline). After each turn we asked the participants to rate both robots' likability individually. Our results show that the participants liked a robot with a positively attributed form of humor significantly more than its gloating robotic interaction partner. However, likability ratings showed a trend to approach each other when either robot laughed or when both robots laughed together. Both, the higher likability ratings for a robot showing positively attributed humor and the decreasing difference in likability ratings when both robots laugh together, provide proof of the positive effect of humor. While participants' age did not affect likability ratings, there was a significant interaction effect between participants' gender and robot type. Female participants rated the iCat more likable, while male participants liked NAO better. In addition, more neurotic people liked the self-ironic robot more when no robot laughed and more open people like the robot showing Schadenfreude more when both robots laughed.

I. INTRODUCTION

Robot personality traits, as a means for enhancing the interaction between humans and robots, are mostly complex concepts which are difficult to grasp. It is difficult to transfer such intricate notions known from interpersonal communication onto human-robot interaction (HRI). Humor is one such concept and possibly the most complex cognitive function in the entire animal kingdom [2]. There are many sides to humor that can hardly be gathered in one picture. We assume that this complexity, on the other hand, makes humor a powerful tool in HRI.

As a first step analyzing humor, we hypothesized that robots expressing different kinds of laughter, score differently in their likability ratings. We performed a 2 (self-ironic laughter vs. no laughter) x 2 (Schadenfreude vs. no laughter) within-subjects user study in which participants watched an iCat and a NAO robot interact with each other. After each of

Nicole Mirnig, Susanne Stadler, Gerald Stollnberger, Manuel Giuliani and Manfred Tscheligi are with Center for Human-Computer Interaction, Christian Doppler Laboratory "Contextual Interfaces", Department of Computer Science, University of Salzburg, Sigmund-Haffner-Gasse 18, 5020 Salzburg, Austria nicole.mirnig@sbg.ac.at

[1]The noun *Schadenfreude* is a loanword originating from German and means "delight in another's misfortune" [1]

the four conditions, participants were asked to rate the likability for both robots separately. Our results showed that the behavior of the robots significantly influenced their likability ratings. To ensure that the difference in likability ratings was not due to the difference in the robots' appearance, we performed a follow-up survey in which robot likability was assessed only in terms of appearance. The results of the survey did not show significant differences between the likability ratings and, thus, confirmed that the behavior of our robots overruled their appearance. Our results can be used to direct future research in exploring robot humor to more detail and to create rounder robot characters.

II. BACKGROUND

Humor is a concept dating back as far as ancient Greece [2]. It is well-researched in fields such as psychology and sociology and lately receives increasing interest in the area of human-computer interaction (HCI). Studies have been reported where humor was attributed the ability to make HCI more natural and flexible (e.g., [3]). Humor has also been associated with a series of positive attributes, such as social competence, intimacy, trust, and stress reduction. For an overview on related work, refer to Hampes [4].

Upon investigating the potential effect of humor on the rating of a virtual agent's cooperative abilities, Kulms et al. found that participants rated a funny virtual agent as less cooperative than a neutral one [5]. The researchers hypothesized that this might be due to the participants' impression of the agent not taking the task seriously enough. Interestingly, in the same study, humor showed to facilitate the interaction in terms of enjoyment and rapport.

Mulder and Nijholt found that humor has the ability to make a computer seem more human when it fails, which, in consequence, can ease the interaction [3]. The researchers also assumed self-depreciating humor an appropriate form to be used for more social HCI [3]. Similarly, Knight found that robots which acknowledge their failure and add a self-deprecating joke are likely to induce more positive feelings in their human interaction partners [6]. Devillers et al. state that humor can be used as a means to establish social relationships and, even more targeted, to recover from faulty situations [7]. Lee et al. found a robot showing mitigation strategies for error recovery (i.e., apology or offer compensation) to create a better impression within a human interaction partner than a robot without recovery behavior [8]. The authors of this study claim that tailored recovery strategies would be even more beneficial as they include the

individuals' personal preferences and expectations. Along the same line, we assume that humor could be a means to recover from robot-induced errors in HRI.

Several studies could already show the positive effect of robot humor on the quality of HRI. Huang and Szafir found a positive effect of humorous behavior on the likability of both, a human and a robotic tutor [9]. Niculescu et al. found that a robot exhibiting humorous behavior was rated more appealing and the tasks to be completed with it were rated more enjoyable. The robot was also rated as more extrovert and was attributed a stronger personality than a robot without humorous behavior [10]. In an experiment on robot humor by Sjöbergh and Araki, jokes were rated funnier when told by a robot as compared to a written form of the jokes. The researchers also explored if the behavior of a second, co-present robot which either laughed along or booed, influences the funniness rating. They found that jokes were rated significantly funnier when the co-present robot laughed or booed than when it showed no reaction [11]. Nonverbal robot humor expressed through movements, sounds, and gestures was found to be rated more humanlike, more likable, funnier, and more entertaining in a study by Wendt and Berg [12]. Katevas et al. [13] found that a robot that gazed at audience members when performing stand-up comedy, was able to elicit more positive audience responses. The researchers could show a tendency that gestures have a positive effect likewise.

Humor and laughter are related, but they are not synonyms. While humor denotes the underlying cognitive process, laughter is an activity that is often triggered by experiencing humorous stimuli [2]. It is possible to correlate different kinds of laughter with positive or negative emotions. Devillers and Soury state that automatic laughter detection and classification of the human behavior could be used to steer a robot's reactions accordingly [14]. Affective and social dimensions are essential in building communication systems with social skills including humor and other informal socially oriented behaviors. Laughter is one such aspect that is being researched with the goal of establishing an intelligent user interface with social communication skills [7]. Finally, "Humoroids" could be a new class of talking agents which are able to express humorous behavior [15].

With our research, we aimed at extending the state of the art on social robots with knowledge around robots that exhibit self-irony and Schadenfreude. We believe that certain kinds of laughter can be specifically implemented to achieve designated functions in HRI and we aimed at exploring how these kinds of laughter could work with robots. Our focus thereby was on the interactional aspects and effect of laughter and not on algorithms and computational details.

III. METHOD

We performed a user study to explore the effect of laughter in HRI. We subjected our participants to two different kinds of robot laughter, that is self-irony and Schadenfreude, and assessed if the kind of laughter affects how participants rate the likability of our robots. We performed a follow-up survey

to ensure that the appearance of our robots did not bias our user study results.

A. User Study Design

Our user study was set up within subjects, with a 2x2 design. The participants watched an interaction between a stationary iCat and a mobile NAO robot (see Fig. 1). The feasibility of robot-robot interaction studies for exploring social robotics was substantiated by Kroos and Herath [16], and the feasibility of using NAO and the iCat together in a study was provided by Cohen et al. [17].
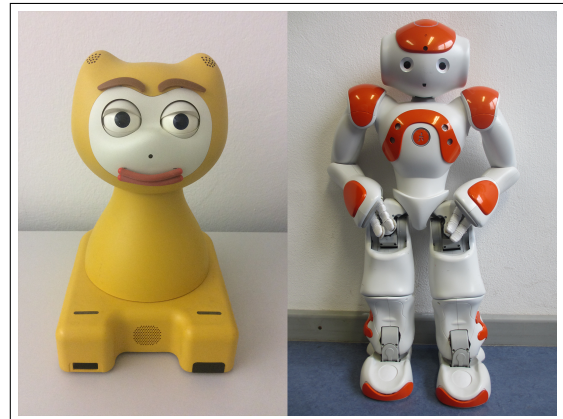

Fig. 1.   iCat (left) and NAO (right).

The robots acted out the following interaction: NAO walked past the iCat, verbally teasing it. The iCat's verbal utterances were neutral. Only a few steps later, NAO stumbled over an obstacle (see Fig. 2). Thereby, NAO either laughed or did not laugh at itself (self-irony) and the iCat either laughed or did not laugh at NAO (Schadenfreude). Superiority theory [3] provides a theoretical explanations for this reversing of the winner and loser roles. It is the falling event that reverses the roles between the robots that creates the humor element in our scenario.

The user study was performed in Austria with all materials in German. The dialog of the scenario stayed the same between conditions, only the laughter setup was manipulated. The following dialog protocol is a translation of the original conversation between the robots.

NAO: *(passing iCat)* "And who are you?"

iCat: *(looks puzzled)* "My name is iCat and I was invented by Philips in 2005."

NAO: "Ha! I am one of the best, most expensive, and furthest developed human-like robots."

iCat: *(looks sad)* "Oh well, at least I can express emotions with my face!" *(looks proud)*

NAO: "You are not able to hold a candle to me, you rusty bucket."

iCat: *(looks sad)*

NAO: "You can't even walk and you're dusty and old as the hills." *(NAO turns its back on the iCat and walks. After a few steps, NAO stumbles and falls.)*

In our setup, NAO laughed at itself (self-irony) and the iCat laughed at its robot interaction partner NAO (Schaden-
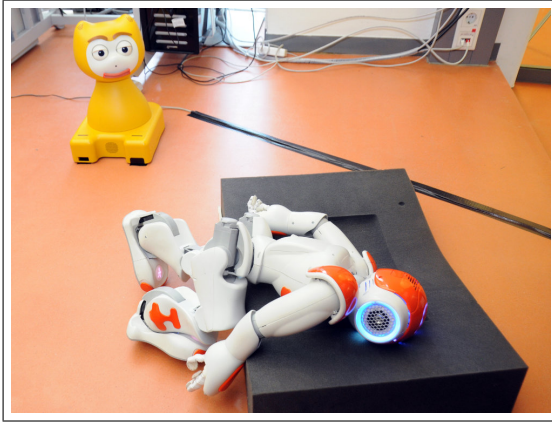
Fig. 2. NAO falling onto the mat while walking away from the iCat.

a short briefing about the experiment. After reviewing and signing the consent form, the participants were asked to fill in a short questionnaire covering basic demographics, their pre-experience with robots, and their general affinity for technology. Then, the participants were asked to complete a questionnaire assessing their personality. Next, the study started. After each condition, the participants were asked to rate the likability of both robots. After completing all four conditions, the user study was concluded with a short de-briefing. Fig. 3 shows a participant observing the interaction between NAO and the iCat.
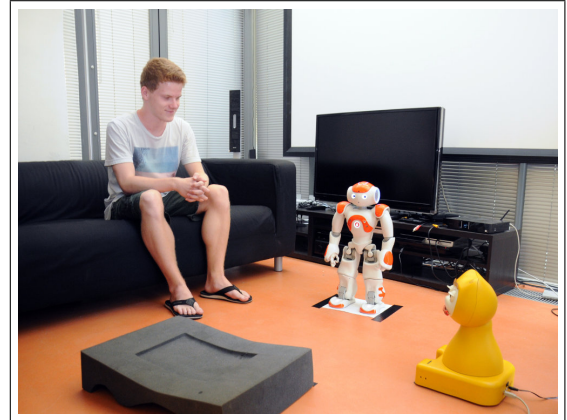

Fig. 3. NAO and the iCat interacting while the participant watches.

1) Dependent Measures: The Big Five Inventory (BFI) Questionnaire ([18]) consists of 44 5-point Likert-scaled items, constructing the five subscales extraversion, agree-ableness, conscientiousness, neuroticism, and openness. With this questionnaire we aimed to explore if differences in our participants' personality affect their ratings of the robots' likability. To assess how much our participants liked NAO and the iCat, the participants were asked to complete the lik-ability scale from the Godspeed Questionnaire Series ([19]). The robots had to be rated individually, meaning that the participants completed the scales twice after each condition, once for NAO, and once for the iCAT. The likability scale consists of 5 items (5-point Likert-scaled).

2) Participants: A total of 22 participants took part in our user study. They were all German native speakers (10 male, 12 female), with a mean age of 28.18 years (SD = 8.10). More than half of our participants indicated having been in contact with robots before (n = 13). Most of those mentions happened in the professional environment (n = 11). Participants' technology affinity was rated above average at 3.86 (SD = .74, 5-point Likert-scaled).

freude). By means of combining laughter with no laughter of both robots, there resulted four conditions (see Table I). Condition A, in which no robot laughed, served as our baseline condition. We have included this baseline with the same verbal and motion behavior than the laughter conditions, to ensure that our results can be traced back to the laughing behavior of the robots. We chose a neutral laughter to ensure that the robots' laughing does not carry uninten-tional connotations. All participants watched all conditions and were asked to rate both robots' likability individually and after each condition. The order of how the participants watched the conditions was randomized to prevent carry-over effects. In our background section we have included related work that shows how nonverbal elements are powerful in eliciting audience reactions. Since verbal humor is difficult to display with robots, we outlined a physical scenario. The affordances of our robots predefined their roles. Since the iCat is a robot head unable to walk, the experiment was not counterbalanced across robot forms.

TABLE I
OVERVIEW ON USER STUDY CONDITIONS

Condition	iCat	NAO
A	-	-
B	laughs (Schadenfreude)	-
C	-	laughs (Self-Irony)
D	laughs (Schadenfreude)	laughs (Self-Irony)

We implemented the utterances of the robots through Choregraphe (NAO) and Animation Editor (iCat). In order to prevent NAO from damage, we placed a soft mat at the intended fall position (see Fig. 2). The setup was approved in our pretest.

B. User Study Procedure

The user study was conducted in the usability lab of the Center for Human-Computer Interaction, University of Salzburg. The participants took part individually, with each person completing all four conditions. The procedure started with the experimenter greeting the participants and providing

C. Follow-up Survey

Following up on the fact that we had deployed two differ-ent robots, we draw on Mori's uncanny valley theory [20]. The theory includes two separate graphs for appearance and behavior. The behavior curve is more dominant, indicating that a robot's behavior overrules its appearance (see Fig. 4). Hoffman provides examples from character animation and

robotics to show that motion can easily overrule appearance [21].

To ensure that any difference in likability ratings was not due to the fact that we used two different robots with unique appearances, we performed a follow-up survey. In this survey, the robots were presented to 22 participants out of context. We placed the two robots next to each other on a table and they did neither move nor speak. We asked the participants to look at the robots and individually rate the likability in terms of the robots' appearance on the Godspeed likability scale (similar to the main user study). We conducted the follow-up survey a couple of weeks after the main user study, with 13 new users and 9 users who had already participated in the main user study. With this, we wanted to achieve a balanced group of participants. With this evaluation we only targeted the robots' appearance. To account for the robots' role and their expressiveness, we had included the baseline condition (A) in our main user study. The robots' laughter was targeted with the experimental conditions (B-D).
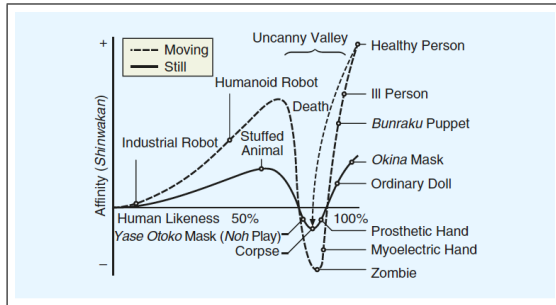
Fig. 4. The Uncanny Valley effect [20].

IV. RESULTS

In the following, we provide the results of our main user study as well as the follow-up survey.

A. Main User Study

We calculated the Godspeed Likability scores for both robots over all conditions. For an overview on the scores, see Table II. Both Table II and Fig 5 nicely show the consistent increase and decease of likability ratings. While the likability for the iCat started out high, they steadily decreased over the conditions. On the contrary, NAO, who started out low in likability, could steadily increase its ratings. Naturally, the robots' verbal and motion behavior influences their likability. This is why NAO's likability is poorest in the baseline condition where no robot laughs, while the additional laughing behavior in the experimental conditions lessens the margin between the robots.

Next, we explored if there are statistically significant differences in participants' likability ratings between the conditions and/or robots. Since the likability ratings variables were normally distributed (see Table III, only one scale slightly deviated from normal distribution, we decided to use a parametric procedure), we performed a mixed factorial ANOVA with gender as the between subjects factor.

TABLE II

GODSPEED LIKABILITY SCALE SCORES PER ROBOT AND CONDITION

Condition	Laugh	iCat (SD)	NAO (SD)
A	-	4.07 (.719)	2.00 (.687)
B	iCat	3.90 (.740)	2.11 (.832)
C	NAO	3.89 (.676)	2.26 (.655)
D	both	3.64 (.612)	2.35 (.682)

TABLE III

OVERVIEW ON RESULTS FOR KOLMOGOROV-SMIRNOV TEST FOR NORMAL DISTRIBUTION PER ROBOT AND CONDITION (ONLY IN CONDITION C, NAO'S RESULT WAS MARGINALLY SIGNIFICANT).

Condition	iCat	NAO
A	$D(22) = .130, p = .200$	$D(22) = .162, p = .139$
B	$D(22) = .124, p = .200$	$D(22) = .133, p = .200$
C	$D(22) = .110, p = .200$	$D(22) = .186, p = .046*$
D	$D(22) = .118, p = .200$	$D(22) = .156, p = .172$

Mauchly's test indicated that the assumption of sphericity had been violated for the main effect of robot behavior, $\chi^2(5) = 15.58$, $p = .008$. Thus, we had to correct the degrees of freedom and used Greenhouse-Geisser estimates of sphericity ($\epsilon = .63$ for the main effect of robot behavior).

There was a significant main effect of robot type on likability ratings. Contrasts revealed that the ratings of the iCat, $F(1, 20) = 103.48$, $p = .000$, $r = .92$, were significantly higher than the ratings of NAO (see Fig. 5). This comparison yielded a very large effect size. In other words, participants rated the iCat more likable than NAO.
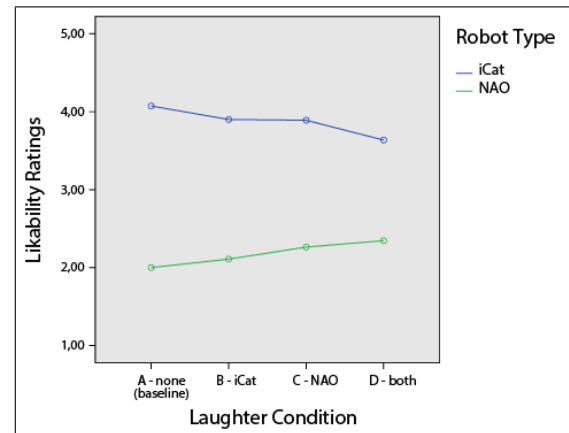
Fig. 5. Effect of robot type and condition on likability ratings.

There was a significant interaction effect between the type of robot and the behavior of the robots, $F(3, 60) = 4.99$, $p = .006$. This indicates that the robots' behaviors had different effects on people's likability ratings. Contrasts were performed to break this interaction down. These revealed a significant interaction effect when comparing the iCat to NAO for condition A (no robot laughs) with D (both robots laugh), $F(1, 20) = 11.40$, $p = .003$, $r = .60$. The statistically significant effects yielded a large effect size.

This means that the margin between NAO's and the iCat's likability differed significantly between condition A (largest margin) and condition D (smallest margin). Therefore, it can be said that robot type and robot behavior combined had a significant effect in both cases, when no robot laughed and when both robots laughed.

There was no significant main effect of gender, indicating that likability ratings from male and female participants were similar, $F(1, 20) = .193$, p = .665, r = .10. However, there was a significant interaction effect between robot type and gender, $F(1, 20) = 6.27$, $p = .021$, r = .49. This effect yielded in a medium effect size. While female participants rated the iCat more likable, male participants rated NAO more likable.

We further explored if participants' age or big five ratings influenced their likability ratings. Therefore, we ran correlations between likability ratings and age/big five scale results. The subscales extraversion, neuroticism, and openness resulted in high reliabilities (Cronbach's α = .79, .78 and .85). Agreeableness and conscientiousness had a relatively low reliability and were, thus, excluded from further analysis (Cronbach's α = .56 and = .60). Participants' age and their ratings of the remaining big five subscales did not deviate significantly from normal (see Table IV). Our sample was in general more extrovert (*Mean* = 3.63 (*SD* = .61) and open (*Mean* = 3.68 (*SD* = .67) and about average neurotic (*Mean* = 2.90 (*SD* = .62).

TABLE IV
OVERVIEW ON RESULTS FOR KOLMOGOROV-SMIRNOV TEST FOR PARTICIPANTS' DEMOGRAPHICS.

Variable	Kolmogorov-Smirnov
Age	$D(21) = .153$, $p = .200$
Extraversion	$D(22) = .101$, $p = .200$
Neuroticism	$D(22) = .164$, $p = .128$
Openness	$D(22) = .109$, $p = .200$

In the following paragraph, we only report Pearson's r for significant differences of the correlations we ran. There was no significant relationship between participants' age and their likability ratings and also not between participants' extroversion and their likability ratings. However, the likability ratings for NAO when no robot laughed were significantly related to participants' neuroticism, $r = .426$, p = .048. This means more neurotic people liked NAO better in the condition when no robot laughed than less neurotic people. The likability ratings of the iCat in the condition when both robots laughed were significantly related to participant's openness, $r = -.439$, p = .041. This means that more open people liked the iCat less when both robots laughed.

B. Follow-up Survey

Our participants' (11 male, 11 female) mean age was 32.36 years (*SD* = 6.71). The likability scale for the iCat resulted in high reliability (Cronbach's α = .90), for NAO in sufficient reliability (Cronbach's α = .70). Next, we calculated the scores for both scales. Both robots' mean ratings were quite

similar (iCat: *Mean* = 3.54 *SD* = .94; NAO: *Mean* = 3.58 *SD* = .46). Both likability scales were normally distributed (iCat: $D(22) = .131$, $p = .200$); NAO: $D(22) = .131$, $p = .200$)). Following up on the differences in likability ratings between the genders which we found in our main study, we conducted a factorial repeated-measures ANOVA. Sphericity was not an issue since we had only two 'conditions' (= robot types in the follow-up survey). We neither found a significant effect for robot type ($F(1, 20) = .037$, $p = .849$), nor a significant interaction effect between robot type and participants' gender ($F(1, 20) = .120$, $p = .733$). This means that there was no difference in likability ratings for the two robots, neither in general, nor between the genders.

V. DISCUSSION

Our results indicate that participants liked the iCat generally better than NAO. This makes sense, since the iCat exhibited a more neutral behavior. It did not mock NAO or respond to NAO's insults and it only expressed Schadenfreude after NAO fell.

Striking is the fact that the difference between the robots' likability ratings is biggest when no robot laughs (condition A) and smallest when both robots laugh (condition D). The likability ratings of each robot taken individually, do not differ significantly between conditions. However, the likability ratings of NAO and the iCat approach each other the more positive the scenario becomes (see Fig. 5).

Laughing at someone is likely not seen as a positive deed. In condition C, however, a more positively attributed form of laughter was shown by NAO, who laughed at itself when it fell. In condition D, finally, both robots laughed. This approximation of likability ratings indicates that both forms of laughter, self-irony and Schadenfreude, carry potential to influence a robot's likability. In this spirit, laughter could be seen as a means to ease a negative situation. For example, when a robot is trying to recover from an error in an interaction. After his teasing verbal behavior towards the iCat, NAO laughed at itself when it fell. The potential herein is the fact that someone capable of laughing at himself, might be perceived less negatively. Irony and self-irony are an important concept of personality, for example, especially in the British culture [22]. Given that we found an impact of self-irony in HRI among German-speaking participants, it can be assumed that irony would be especially powerful in a British context.

We attribute the differences in likability ratings to both robots' laughing behavior, since the dialog was not varied between conditions. However, the role an interaction partner inhibits is important. While the iCat was more neutral, NAO behaved more affronting. In this study we have not measured how our participants perceived the roles of the robots. This is an interesting point to be researched in the future.

We could additionally support our results with our follow-up survey. The data on robot likability rated out of context, showed no difference between the iCat and NAO. This result goes in line with Mori's uncanny valley in which it is stated that a robot's behavior overrules its appearance. In

our study, we could accordingly attribute the difference in likability ratings to the different laughing behaviors of the robots, which overruled their appearance. It is interesting to notice that both robots break their character. The neutral dialog of the iCat is followed by Schadenfreude, while the teasing dialog of NAO is followed by self-ironic laughter. This can be taken as an extension of the uncanny valley. While the role of a robot seems most important in terms of a robot's likability, laughter seems to have an influence too, followed by appearance which is overruled by both, the role and laughter.

In our main user study, we found an interaction effect between gender and robot behavior. Female participants rated the iCat more likable than NAO, and male participants rated NAO more likable than the iCat. With the results of our follow-up survey, where we did not find a statistically significant difference in likability ratings between the genders, we could rule out that this gender-difference was rooted in the fact that we had used two different robots. Since we found a difference in the user study where the robots acted within a scenario and we found no difference when only the robots' outer appearance was rated, this result provides further evidence that behavior overrules appearance.

Neither participants' age nor their level of extroversion significantly influenced their ratings of our robots' likability. However, neurotic people preferred NAO in the condition when no robot laughed and more open people liked the iCat less when both robots laughed.

We have stated previously that humor and laughter are closely related, but they are not the same. Humor is a cognitive process and laughter an activity that is often triggered by a humorous experience. Laughter is an activity that we have used specifically to manipulate our participants' perception.

VI. CONCLUSIONS

We believe in the potential of humor and laughter in HRI. In our study, we found significant differences in the likability ratings between the two robots, with our participants preferring the robot showing a more neutral verbal behavior, even though it laughed about its robotic interaction partner. However, we found that laughter, in general, holds potential to influence the quality of how humans perceive robots. For future research, we take away that different forms of laughter may be used as a means to make the interaction between humans and robots more natural and possibly more enjoyable. Furthermore, we could show that laughter can be used to improve how a robot with an unpleasant role is perceived. This could be especially relevant for robot developers and HRI researchers who aim at creating more believable robot characters.

ACKNOWLEDGMENT

REFERENCES

[1] Schadenfreude, "Merriam-Webster's Learner's Dictionary," May 2016. [Online]. Available: http://www.merriam-webster.com/dictionary/schadenfreude

[2] J. Polimeni and J. P. Reiss, "The first joke: exploring the evolutionary origins of humor," *Evolutionary Psychology*, vol. 4, pp. 347–366, 2006.

[3] M. P. Mulder and A. Nijholt, "Humour research: State of art," Tech. Rep., 2002.

[4] W. P. Hampes, "The relation between humor styles and empathy," *Europe's Journal of Psychology*, vol. 6, no. 3, pp. 34–45, 2010.

[5] P. Kulms, S. Kopp, and N. C. Krämer, "Let's be serious and have a laugh: Can humor support cooperation with a virtual agent?" in *Intelligent Virtual Agents*. Springer, 2014, pp. 250–259.

[6] H. Knight, "Eight lessons learned about non-verbal interactions through robot theater," in *Social Robotics*. Springer, 2011, pp. 42–51.

[7] L. Devillers, S. Rosset, G. D. Duplessis, M. A. Sehili, L. Béchade, A. Delaborde, C. Gossart, V. Letard, F. Yang, Y. Yemez *et al.*, "Multimodal data collection of human-robot humorous interactions in the joker project," in *ACII 2015*. IEEE, 2015, pp. 348–354.

[8] M. K. Lee, S. Kiesler, J. Forlizzi, S. Srinivasa, and P. Rybski, "Gracefully mitigating breakdowns in robotic services," in *HRI 2010*. IEEE, 2010, pp. 203–210.

[9] C. Huan and D. Szafir, "No joke: examining the use of humor in computer-mediated learning," 2001.

[10] A. Niculescu, B. van Dijk, A. Nijholt, H. Li, and S. L. See, "Making social robots more attractive: the effects of voice pitch, humor and empathy," *Int. J. of Social Robotics*, vol. 5, no. 2, pp. 171–191, 2013.

[11] J. Sjöbergh and K. Araki, "Robots make things funnier," in *New Frontiers in Artificial Intelligence*. Springer, 2009, pp. 306–313.

[12] C. S. Wendt and G. Berg, "Nonverbal humor as a new dimension of hri," in *RO-MAN 2009*. IEEE, 2009, pp. 183–188.

[13] K. Katevas, P. G. Healey, and M. T. Harris, "Robot comedy lab: experimenting with the social dynamics of live performance," *Frontiers in psychology*, vol. 6, 2015.

[14] L. Y. Devillers and M. Soury, "A social interaction system for studying humor with the robot nao," in *ICMI2013*. ACM, 2013, pp. 313–314.

[15] P. Dybala, M. Ptaszynski, R. Rzepka, and K. Araki, "Humoroids: conversational agents that induce positive emotions with humor," in *AAMAS2009*. International Foundation for Autonomous Agents and Multiagent Systems, 2009, pp. 1171–1172.

[16] C. Kroos and D. C. Herath, "We, robots: Correlated behaviour as observed by humans," in *Social Robotics*. Springer, 2014, pp. 229–238.

[17] I. Cohen, R. Looije, and M. A. Neerincx, "Child's perception of robot's emotions: effects of platform, context and experience," *International Journal of Social Robotics*, vol. 6, no. 4, pp. 507–518, 2014.

[18] O. P. John, L. P. Naumann, and C. J. Soto, "Paradigm shift to the integrative big five trait taxonomy," *Handbook of personality: Theory and research*, vol. 3, pp. 114–158, 2008.

[19] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," *International journal of social robotics*, vol. 1, no. 1, pp. 71–81, 2009.

[20] M. Mori, K. F. MacDorman, and N. Kageki, "The uncanny valley [from the field]," *Robotics & Automation Magazine, IEEE*, vol. 19, no. 2, pp. 98–100, 2012.

[21] G. Hoffman and W. Ju, "Designing robots with movement in mind," *Journal of Human-Robot Interaction*, vol. 3, no. 1, pp. 89–122, 2014.

[22] J. Brassett, "British irony, global justice: a pragmatic reading of chris brown, banksy and ricky gervais," *Review of International Studies*, vol. 35, no. 1, pp. 219–245, 2009.