Effects of a Robotic Storyteller's Moody Gestures on Storytelling Perception

Junchao Xu, Joost Broekens, Koen Hindriks
[1]Department of Intelligent Systems
Delft University of Technology
Delft, the Netherlands
junchaoxu86@gmail.com, {k.v.hindriks, d.j.broekens}@tudelft.nl

Mark A. Neerincx[1,2]
[2]Human Factors
TNO
Soesterberg, the Netherlands
mark.neerincx@tno.nl

***Abstract*—A parameterized behavior model was developed for robots to show mood during task execution. In this study, we applied the model to the coverbal gestures of a robotic storyteller. This study investigated whether parameterized mood expression can 1) show mood that is changing over time; 2) reinforce affect communication when other modalities exist; 3) influence the mood induction process of the story; and 4) improve listeners' ratings of the storytelling experience and the robotic storyteller. We modulated the gestures to show either a congruent or an incongruent mood with the story mood. Results show that it is feasible to use parameterized coverbal gestures to express mood evolving over time and that participants can distinguish whether the mood expressed by the gestures is congruent or incongruent with the story mood. In terms of effects on participants we found that mood-modulated gestures (a) influence participants' mood, and (b) influence participants' ratings of the storytelling experience and the robotic storyteller.**

Keywords—Storytelling; Mood Expression; Body Language; Social Robots; Human Robot Interaction.

I. Introduction

Bodily expression is important for social robots to naturally communicate affect to humans [1]. Expressive body language of a robot facilitates human understanding of a robot's behavior, rationale, and motives [2], and increases the efficiency of human-robot task performance and robustness [3]. It is known to increase the perception of a robot as trustworthy, reliable, and life-like [4]. Bodily affective expression is in particular important for humanoid robots that lack facial features such as NAO, ASIMO, and QRIO. A model that enables robots to show mood during tasks has been developed [5]. In this study, we applied this model to coverbal gestures of a storytelling scenario to study the mood expression. We have several motivations.

First, in previous studies mood expression based on parameterized behavior has been set to show mood at fixed discrete levels, e.g., a positive and a negative level. It is not clear whether this model is able to show mood that is changing continuously over time. In this study, we apply the model to the coverbal gestures of a storytelling scenario, and we modulate the parameters of coverbal gestures continuously, in order to show an evolving mood that is congruent with the story mood changing with the story line. We evaluate whether people perceive the body language as changing over time and congruent with the story line.

Second, in naturalistic settings people perceive affective information from different channels simultaneously. Interactions between bodily expression and other modalities of expression have been found in people's perception [6], [7], [8], [9], [10]. We would like to investigate the use of the affective body language in a scenario where a robot also communicates affect through other affective channels. We need to guarantee that the robot's bodily expression generated by our model can express a congruent mood with other modalities, and we expect that the introduction of bodily expression can improve the recognition and the effects of the overall expression. Speech is an inherent channel of affective communication in storytelling. In this study, we report our exploration of the interaction between body language and speech semantics, while we kept the voice features the same.

Third, we'd like to see whether the affective body language is able to improve storytelling experience. Storytelling is an important application, for example, it supports children's development [11]. Improving storytelling experience may increase acceptance of the robot application. Finally, using the mood expression model in a storytelling context provides more evidence about whether the model is generalizable to different applications.

II. Related Work

Bodily expression was found to influence or be influenced by other modalities of expressions. Stock et al. [6] found that bodily expression influenced the recognition of facial expression and emotional tone of a voice. Later, Stock et al. also found that recognition of bodily expression was influenced by nonverbal auditory information (human and animal sounds) [7] or task irrelevant auditory (music) [8]. Meeren et al. [9] studied people's perception of congruent and incongruent integration of facial expression and emotional body language using photographs. They found that people's judgment of facial expression was biased towards the bodily expression. Kret et al. [10] showed that congruency between facial and bodily expression improved recognition. In our case, we explore how the affective body language of robots interacts with affect expressed by semantics of stories.

Gestures influence people's perception of the communication quality between robots and people. Salem et al. studied how gesture influences humans' evaluation of communication quality and the robot using ASIMO [12], [13]. Results showed that the robot was rated more positively when coverbal gestures were used compared to speech alone, even when the gestures did not semantically match the speech. An interesting result is that incongruent gestures were even rated higher across many aspects. Their explanation is that in the incongruent condition the robot is less predictable. In our study, we also include a congruent and incongruent gesture condition, but we focused on affective

978-1-4799-9953-8/15/$31.00 ©2015 IEEE

IEEE computer society

congruency, i.e., whether the robot gestures show mood that is congruent with story content. In this study gestures are manually coordinated with speech. Automatic coordination is beyond the scope of this study (for an overview see [14]). Gaze of a robot during storytelling is important [15]. In our study, the robot always looks at the listener when the robot does not perform head movement.

Emotional coverbal gestures for storytelling were usually built based on corpora of human behaviors. For example, expressive coverbal gestures of a NAO robot used for a storytelling scenario were constructed using a video corpus of human storytellers [16]. The gestures were shown to improve participants' perception of the expressivity of the robot storyteller [17]. Park et al. developed an expressive robot behavior generation framework based on sentence types and emotions [18]. The behaviors were generated based on movements of actors. In this study, the affective gestures are generated by a parameterized behavior model (Fig. 1). The principles of parameter modulation were obtained from users [5] and the resulting gestures have been evaluated in [19], [20].

Robot storytelling is an important application. For example, it can be used for children education. Montemayor et al. [21] provided children with a tool to create robotic pets and stories, which then can be acted out by the robots. Emotion expression was argued to be an indispensable feature of a robotic storyteller, as children typically attribute emotions to toys they play with. Storytelling was used as an educational activity to test whether the KindSAR robot can engage children in constructive learning [11]. Bodily expression, alongside with facial expression and vocal expression, was used to show the robot emotion, for example, happiness was shown by raising hands, nodding head, and eye light blinking. Results indicated that the story emotion was efficiently conveyed by the robot and children's emotional involvement was promoted, as the children's emotional responses were significantly correlated with the story emotion. I big difference with our study is that we use mood expression that is expressed using parametrized gestures.

III. Mood Expression In Storytelling

Our work focuses on mood. Distinctions between affect, emotion, and mood are explained in [22], [23], [24]. Here, we highlight the distinctions between mood and emotion that are related to expression: emotion is a short-term, intense affective state, associated with specific expressive behaviors; mood is a long-term, diffuse affective state, without such specific behaviors. Mood emphasizes a stable affective context, while emotion emphasizes affective responses to events. We use valence and arousal dimensions to represent mood.

We have used a parameterized behavior model (Fig. 1) for integrating affect expression with functional behaviors (e.g., task behaviors, communicative gestures, and walking). Using this model, robot movements can be modulated to display the robot mood by changing behavior parameters with respect to both spatial extent and motion dynamics [5], [25]. This model enables a robot to express mood, even during task execution by modulation of the "style" of the behavior. The resulting mood expressions have been evaluated with the NAO in a laboratory setup without context [19] and in a game setting [26].

This model has been applied to 41 coverbal gestures and used in a university lecture scenario [20]. We reused these gestures in the storytelling scenario, in order to express the story mood while the robot is telling stories. The gestures were manually selected for the sentences of the stories and manually aligned with the words

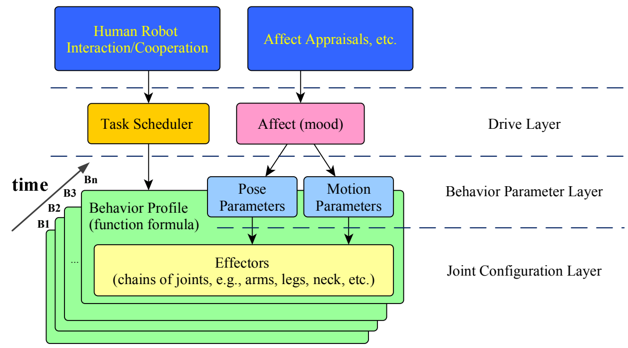

Fig. 1. General Parameterized Behavior Model.

in the sentences. In study [20], a script engine was designed to orchestrate the robot gestures, speech, and slides. We reused this engine in the storytelling study. The robot speech was generated by a Text To Speech engine shipped with the robot. The script engine synchronizes the starting points of texts and its coverbal gestures automatically. The robot selected leg movements randomly from a predefined corpus in real time and performed them at the same time as hand gestures. Random leg movements are used to maintain a life-like quality of the robot. Pilot testing showed that a talking robot standing still was perceived unnatural.

IV. Questions and Hypotheses

Because the parameters of the gestures are controlled by a continuous variable and can be modulated in real time, the gestures can be modulated to show a continuously changing mood. This study first investigates if the parameterized behavior model can be used to generate behavior that expresses a mood that *changes over time*. Because story mood changes throughout a story, we chose the storytelling domain. We chose stories in which the mood expressed semantically changes over time. We reasoned that if listeners perceive the mood expressed by the robot as congruent with the story mood, it must have been following the story mood over time as the story mood changes over time. To test this, we hypothesize the following:

H1. When robot mood is congruent with the story mood, listeners rate the congruency of the robot mood with the story mood to be higher, as compared to a robot mood that is the opposite.

Second, we investigate what perceived effects affective robot body language has on storytelling. The affective communication in storytelling is inherently a multimodal communication, since affect is conveyed through 1) the semantics, i.e., the story content; 2) the voice; and 3) the body language. Body language of humans was shown to influence the recognition of emotions from other modalities [6], [7], [8], [9], [10]. We would like to see if robot body language has similar effects on mood recognition. The difference is that we investigate the effect of robot body language on the affect conveyed by semantics. Here we modulate the robot body language depending on the story mood, and we do not manipulate the robot voice. It was shown that body language reinforced people's recognition of the robot emotions on top of facial expressions [27], [28]. Body language thus may also be able to reinforce other forms of expression, e.g., affect expressed in stories. Specifically, we are interested in whether robot body language can facilitate the understanding of the story mood and make the story mood perceived stronger. We test these aspects based on listeners' self-reports:

H2. A) When robot mood is congruent with the story mood, listeners perceived the body language as helpful in understanding the story mood, as compared to the incongruent condition.

B) When robot mood is congruent with the story mood, listeners perceive that the body language makes the story mood stronger, as compared to the incongruent condition.

Third, it is known that stories can induce emotions or moods to listeners [29]. Further, it is well known that mood can be transferred from one person to another [30]. Previous studies also showed that mood can be transferred from a virtual agent displaying facial expressions [31] or a robot displaying affective body language [20], [26] to a person that is interacting with the agent/robot. Body language provides a second channel of mood induction. Moreover, if the perceived story mood is reinforced (H2 and H3) the mood induction may also be stronger. We thus hypothesize:

H3. When robot mood is congruent with the story mood, listeners report a stronger mood change for their own mood, compared to the incongruent condition.

Finally, it was found in a university lecture study [20] that affective body language was able to influence students' ratings of the robot. Creed and Beale [32] found that inconsistent displays of emotion negatively influenced the perception of an embodied agent. Berry et al. [33] also found that the consistency of the emotion expressions influenced the ratings of the virtual agent. We test the following hypothesis:

H4. When robot mood is congruent with the story mood it improves listeners' ratings of the storytelling experience and the robotic storyteller, compared to the incongruent mood condition.

V. Experimental Setup

A. Experimental Design

To test the hypotheses, we defined three conditions:

1. **Congruent condition**: coverbal gestures are modulated to express mood congruent with the mood of the current sentence. The robot also performs random leg movements.
2. **Incongruent condition**: coverbal gestures are modulated to express mood opposite to the mood of the current sentence. The robot also performs random leg movements.
3. **Control condition**: the robot performs no coverbal gestures, but random leg movements.

The control condition provides a benchmark, to check for generic effects of gestures. To rule out the possibility that the hypotheses can be verified by arbitrary modulation, we use the incongruent condition as contrast (a stronger control than random gestures).

We used a between-subject design. Each participant listened to the stories in only one body language condition (RBL condition for short). The dependent variables are 1) the perceived congruency of the coverbal gestures with the story mood; 2) the perceived helpfulness of the coverbal gestures in the understanding of the story mood; 3) the perceived reinforcement of the coverbal gestures on the story mood; 4) listeners' mood; and 5) general ratings of the robotic storyteller.

B. Materials

Two inspiring stories, one realistic (the ice cream story) and one fantasy (the cracked pot story), were chosen for this study. Both stories were taken from this website[1] and modified. The ice cream story lasts for about 1 minute 45 seconds on our system, and the cracked pot story 3 minutes. The full texts of the stories and the videos of the storytelling can be found in the supplementary materials and our web site[2]. Each story had 2 break points (explained later).

To avoid a ceiling effect (i.e., the mood expressed by the story content is already very strong, so the mood added by the body language is limited), we chose stories with moderate mood or emotions. To avoid confusing mood of different characters in the story, we made the narrative focused on one character. The mood (valence) of the story was annotated sentence-wise by five experienced annotators beforehand. Their annotations are consistent: for the ice cream story *Cronbach's* $\alpha = 0.736$; for the cracked pot story *Cronbach's* $\alpha = 0.890$. This annotation was used to drive the gesture-based mood model.

A grey NAO robot (NaoQi version 1.14; head version 4.0; body version 3.3; 58cm tall) with LED lights switched off was used as the storyteller. The robot stands on a table while telling stories and listeners sit in front of the robot while listening.

C. Measures

We test H1, H2A, H2B, and H4 with the following 11 item post-experiment questionnaire. Each question is measured with a statement to be answered on a 5-point (-2 to 2) Likert scale:

1) You did not notice that the robot was performing gestures while it was telling the stories.
2) The robot teller was using gestures to communicate the story mood.
3) The mood expressed by the robot gestures is congruent with the story mood.
4) The gestures of the robot teller helped you to capture the story mood.
5) You mainly captured the story mood from the robot speech.
6) The gestures of the robot teller made the story mood stronger.
7) The robot teller kept you immersed in the stories.
8) The robot teller enthusiastically presented the stories.
9) The robot teller organized the speech and gestures in a fluent way.
10) The gestures of the robot teller are natural.
11) Overall, you are satisfied with the performance of the robot.

Questions Q1 and Q2 check if listeners notice the gestures and realize the gestures are used to communicate mood. Q3 tests H1, Q4 and Q5 test H2A, and Q6 tests H2B. Finally, Q7~Q11 test H4.

The change in listeners' own mood (H3) is measured using the valence and arousal dimensions of SAM (self-assessment manikin) [34] on a 5-point Likert scale. In addition, we asked participants to annotate the story mood during the storytelling in real time in order to get more objective measures of the effects of the modulated gestures on the perception of the story mood (related to H2 and H3). Participants were asked to click the AffectButton whenever they thought the story mood changed.

D. Participants

66 participants (42 males and 24 females) aged 19 to 48 (*Mean* = 28.0, *SD* = 4.8) were recruited from the university campus. They were from 19 different countries: 17 are Dutch; 19 are Chinese;

[1] http://rishikajain.com

[2] http://ii.tudelft.nl/SocioCognitiveRobotics/index.php/Storytelling

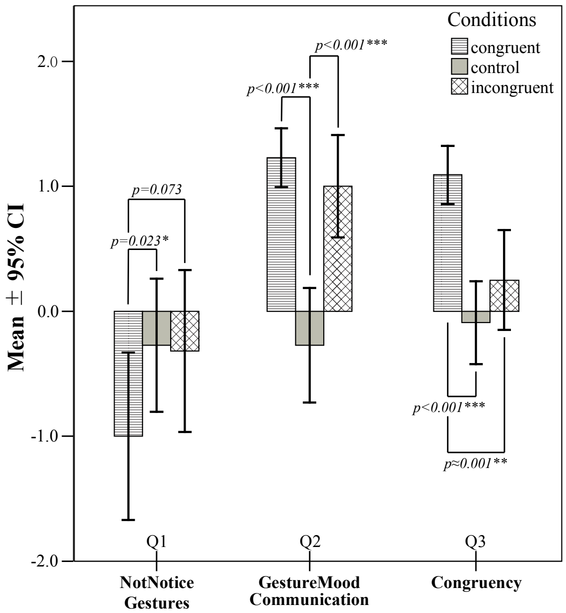

Fig. 2. Perception check and perceived congruency between the robot body language and the story mood

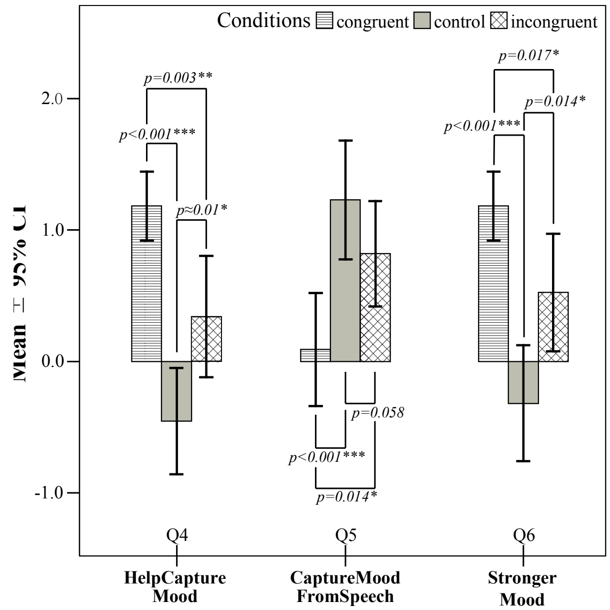

Fig. 3. Effects of the robot body language on the perception of the story mood

and 7 are Indian. A pre-experiment questionnaire confirms that the participants had little expertise on robotics or virtual agents. Participants had some storytelling experience and they held a positive attitude to reading or listening to stories. Each participant received a gift after the experiment.

E. Procedure

Each participant read the experiment instructions, filled out a consent form, and a general questionnaire about demographics and previous experiences with robots and virtual agents. Participants were told to pay attention to the robot in general when the robot was telling the stories, but we did not emphasize mood or behavior to try to eliminate a demand effect (participants rating what they think we want them to feel / see). Then, a training session of the AffectButton started. The task was to adjust the facial expression on the button to match 32 given affective terms [35]. Just before the start of each story (T0), the current mood of each participant was measured with a SAM self-report. When the robot stopped at break point (T1 and T2) during the storytelling or the end of each story (T3), the mood of the participant was also measured using SAM questionnaires. Participants also filled out a questionnaire about whether they understood the story, whether they heard the story before, the perceived story length in minutes, and their attention distribution at T3 after the mood measurement. Participants were allowed to take a break between the two stories. After the two stories, participants filled out the post-experiment questionnaire. After the experiment, participants were debriefed and thanked for participation. The experiment took 30 minutes.

VI. RESULTS

A. Expressing Evolving Mood (H1)

We first check whether participants noticed the robot gestures (Q1) and thought the gestured were used for communication (Q2). Then we test the perceived congruency of the robot gestures (Q3). Kruskal–Wallis tests show a marginal significance for Q1: $\chi^2(2) = 5.568, p=0.060$ and significance for Q2: $\chi^2(2) = 23.447, p<0.001$ and for Q3: $\chi^2(2) = 22.675, p<0.001$. Fig. 2 shows the means and significances of the post-hoc Mann-Whitney U tests.

These results suggest three things. First, participants noticed gestures more in the congruent condition than in the control condition but not in the incongruent condition. This is a bit odd as participants do indicate that the robot uses gestures to communicate mood in the incongruent condition. Second, participants considered that the robot was using gestures to communicate the story mood in both congruent and incongruent conditions and no significant difference between the two conditions was observed. Apparently gestures made the robot more expressive in general. Third, participants perceived the mood expressed by the gestures in the congruent condition as significantly more congruent than in the incongruent condition or in the control condition. As the story mood changes over time (confirmed by pre-experiment annotation), this confirms the model's ability to express mood that is evolving over time.

In sum, these results support our hypothesis that modulated coverbal gestures can be used for communicating story mood in storytelling continuously and that participants are able to distinguish whether the robot coverbal gestures are congruent or not with the story mood (H1). Further, this confirms the importance of congruent gestures as opposed to incongruent.

B. Reinforcement of story mood (H2AB)

Participants' Perception

To test if modulated gestures helped participants to capture mood from the story (Q4), provided an efficient way to capture the story mood in addition to the speech (Q5), and made the perceived story mood stronger (Q6), we performed Kruskal–Wallis tests. These show significant differences for Q4: $\chi^2(2) = 26.979, p<0.001,$ for Q5: $\chi^2(2) = 15.410, p<0.001$, and for Q6: $\chi^2(2) = 22.188, p<0.001$. See Fig. 3 for means and post-hoc Mann-Whitney U.

These results suggest two things. First, participants considered coverbal gestures to be helpful for capturing the story mood in general, whether the gestures are congruent with the story or not. However, congruent gestures were considered significantly more

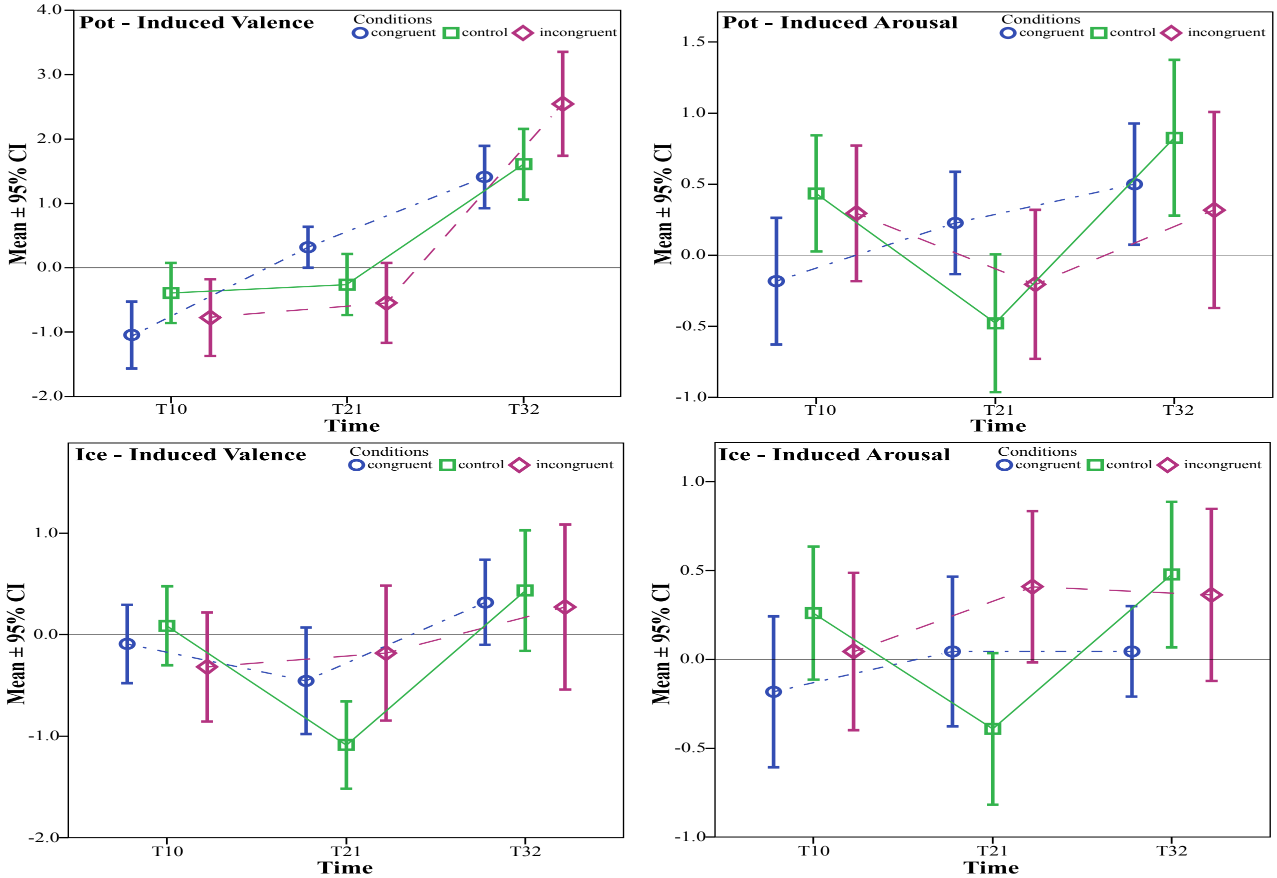

Fig. 4. Induced mood (valence and arousal) in the period T10: from start to break1, T21: from break1 to break2, and T32 from break2 to end. The story texts and the break positions can be found in the supplement and our web (see footnote 2)

helpful (Q4). This is supported by the fact that participants indicated to use the text to capture the story mood in both the control and incongruent conditions (Q5). This indicates that people pay more attention to the mood in the gestures when these are congruent, and otherwise pay more attention to the text. Second, participants indicated that coverbal gestures made the story mood stronger in both congruent and incongruent conditions, with the congruent condition being significantly stronger than the incongruent condition. Apparently, some gesturing is better than none, but mood congruent gestures are better than incongruent gestures.

In sum, we conclude that the affective quality of the gestures influenced these effects: the gestures helped with the understanding the story mood (H2A) and made the story mood stronger to a significantly larger extent, when the gestures expressed mood that is congruent with the story mood (H2B).

Annotation of the Story Mood

To further study mood enhancement by gestures, we test whether the annotated mood is more positive (valence) and more active (arousal) when the mood of the story is positive and active; more negative and more passive when the mood of the story is negative and passive. All participants' annotations of one story are put together and sorted according to time. Then data points were binned into time periods. ANOVA analyses testing for difference between conditions did not reveal significant differences. This indicates the affect traces are similar between conditions.

C. Effects on Participants' Moods (H3)

To find out whether the storytelling induced mood to the participants, we first check whether participants' mood changed over time. A mixed model MANOVA with congruency as between-subject factor and time (T0, T1, T2, and T3) as within-subject factor was used to analyze the participants' mood (valence and arousal). The results of the overall tests show that time has an significant effect on the participants' mood, both for the cracked pot story $F(6,58)=20.242$, $p<0.001$, *partial* $\eta^2=0.677$, and for the ice cream story $F(6,58)=5.371$, $p<0.001$, *partial* $\eta^2=0.357$. This means that participants' moods were influenced by the storytelling. That is, the mood induction occurred. The cause could be the robot body language, the story content, or something.

To test whether robot body language had an effect on the mood induction process, we calculated induced mood in three periods: from T0 (start) to T1 (the first break), from T1 to T2 (the second break), and from T2 to T3 (end) for each story. The changes of the mood variables during the periods were used as induced mood. Fig. 4 illustrates the means and confidence intervals of the induced valence and arousal. A mixed model MANOVA, with congruency as between-subject factor, time as within-subject factor, and valence and arousal as two measures, was used.

For the cracked pot story, the effect of the modulated gestures on the mood induction process is evidenced by a significant interaction effect (time*congruency): multivariate test Pillai's Trace $F(8, 122)=2.300$, $p=0.025$, *partial* $\eta^2=0.131$. The univariate test showed that the interaction effect is on the valence

IEEE computer society

$F(4)=3.193$, $p=0.016$, partial $\eta^2=0.092$. For the ice cream story, an effect of the congruence is observed at a marginal significance level: Pillai's Trace $F(4,126)=2.257$, $p=0.067$, partial $\eta^2=0.067$ (Roy's Largest Root shows significance: $F(2,63)=3.968$, $p=0.024$, partial $\eta^2=0.112$). The univariate test showed that the effect of congruence is on the arousal $F(2)=3.846$, $p=0.027$, partial $\eta^2=0.109$. Post hoc tests with Bonferroni correction showed that the induced arousal in the incongruent condition is significantly larger ($p=0.023$) than in the congruent condition.

Overall, this suggests that robot body language influenced how the participants' mood was evolving over time. Put differently, body language has an effect on the mood induction process. However, a clear effect was shown only for arousal and only for one of the two stories. As such, the results do not support the hypothesis that mood induction was more pronounced in the congruent condition (H3).

D. Experience of the Storyteller (H4)

In this section, we present the results of how the modulated coverbal gestures influence participants' storytelling experience and their ratings of the robotic teller. The results of Kruskal–Wallis tests of Q7~Q11 are listed in TABLE I, and Fig. 5 shows the result of the post hoc Mann-Whitney U tests.

We did not observe a significant effect of gestures on immersion (Q7). Possible explanation is that the story content already made listeners immersed in the context (or not) and that semantic meaning is therefore more important to immersion.

Participants perceived the robot to be significantly more enthusiastic as long as the robot performs gestures, no matter the mood expressed by the gestures is congruent with the story mood or not (Q8). This may be because that more body movements (gestures vs. no gestures) made the robot appear more active in general. In the congruent condition, the perceived enthusiasm is higher than the incongruent condition at a marginal significance level. This means that in general participants considered a storytelling robot that performs coverbal gestures as enthusiastic and the affective quality of the gestures may have influenced participants' perception of enthusiasm.

We did not observe a significant difference in the fluency of gesture-speech organization between congruent and incongruent conditions (Q9). However, congruent gestures were perceived to be more fluently organized with the speech compared to the control (random leg movements). This provides some evidence that mood-congruent modulation of gestures is important for perceived gesture-speech organization.

We did not observe a significant effect of the gestures on naturalness (Q10).

Finally, the overall satisfaction in the congruent condition is higher than incongruent condition at a marginally significant level (Q11). This is in line with the trends on fluency and enthusiasm.

In sum, there is some evidence to support the hypothesis that story-mood-congruent gestures improve storytelling (H4).

VII. General Discussion

Overall, our results seem to indicate that the semantic channel takes priority over robot body language. The attention distribution questionnaire showed that participants paid 26% of attention to the robot movements and 52% to the robot speech.

TABLE I. KRUSKAL-WALLIS TESTS FOR Q7~Q11

Statistics	Q7 Immerse	Q8 Enthus.	Q9 Fluency	Q10 Natural.	Q11 Overall
$\chi^2(2)$	0.449	14.194	6.379	3.119	2.683
sig.	0.799	<0.001	0.041	0.210	0.261

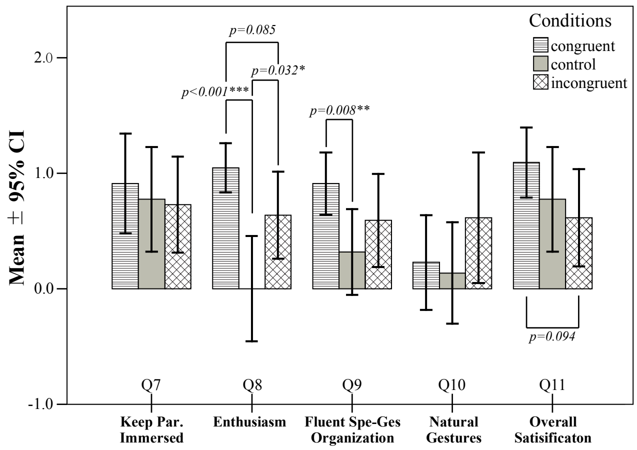

Fig. 5. Ratings of the storytelling experience and the robotic storyteller.

The domination of speech seems to be true especially when speech and body language are incongruent. For example, in the incongruent condition, participants indicated that they did not notice the body language (Q1) and that they only captured the story mood from the spoken text (Q5). This suggests that participants shifted attention to the speech after they perceived little meaning in the incongruent body language. The underlying reason may be that incongruent information adds to cognitive load, which was observed in [33]. The phenomenon that participants following the speech in the incongruent condition can also be explained by cognitive dissonance theory [36], which suggests that people attempt to reduce inconsistency in their perceptions. In our case, participants ignored the mood expressed by the incongruent gestures but followed the mood in the story content.

Several other reasons may also account for the lack of a significant effect of the robot body language on the participants' annotation of the story mood. First, using the AffectButton to annotate the story mood and looking at the robot in real time is difficult. Attention distribution showed that participants paid 22% of attention to the laptop during the storytelling, in addition to the amount of attention already taken by the robot speech. This attention occupation might further reduce the effect of the robot body language. Second, it is methodologically difficult to correctly analyze affect traces over time. For example, we cannot decide with 100% certainty to which bin a measurement belongs as participants were free to rate when they wanted, and rating takes some time. So, some inputs might be wrongly classified to belong to a particular sentence. Last, although the chosen stories are simple, the story mood changes often. This increased complexity and difficulty of annotation. It would be better to do a similar test with a story that clearly changes in mood exactly once.

VIII. Conclusion and Future Work

Our study shows that it is feasible to modulate coverbal gestures in real time, based on the behavior parameterize model, to express a mood that is evolving over time and is congruent with the story line. The results show that participants distinguished whether the mood expressed by the coverbal gestures was congruent with the story mood or not. Results also show that participants perceived the coverbal gestures expressing congruent mood helped them to capture the story mood and made the story mood stronger. In terms of effects on participants we found that mood-modulated robot body language (a) influences participants' mood (but the effect is not entirely clear), and (b) influenced participants' ratings of the storytelling experience and the robotic storyteller.

To the best of our knowledge, this study is one of the very few in which bodily mood expression of robots are studied in a scenario in which another affective communication channel (speech semantics) exists. Some challenges regarding the coexistence of affective robot body language and affective speech content are revealed, which are yet to be explored in the future work. These include *speech content domination of affect*, the *difficulty to rate in real time*, and participants' apparent *reduction of inconsistency* between gesture mood and story mood.

REFERENCES

[1] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," Rob. Auton. Syst., vol. 42, no. 3–4, pp. 143–166, Mar. 2003.

[2] C. Breazeal, A. Takanishi, and T. Kobayashi, "Social Robots that Interact with People," Springer Handb. Robot., pp. 1349–1369, 2008.

[3] C. Breazeal, C. D. Kidd, A. L. Thomaz, G. Hoffman, and M. Berlin, "Effects of nonverbal communication on efficiency and robustness in human-robot teamwork," 2005 IEEE/RSJ Int. Conf. Intell. Robot. Syst., pp. 708–713, 2005.

[4] C. Breazeal, Designing sociable robots. MIT Press, 2002.

[5] J. Xu, J. Broekens, K. Hindriks, and M. A. Neerincx, "Mood expression through parameterized functional behavior of robots," in Proceedings of IEEE International Symposium on Robot and Human Interactive Communication, 2013, pp. 533–540.

[6] J. Van den Stock, R. Righart, B. De Gelder, J. Van den Stock, and B. de Gelder, "Body expressions influence recognition of emotions in the face and voice.," Emotion, vol. 7, no. 3, pp. 487–94, Aug. 2007.

[7] J. Van Den Stock, J. Grèzes, and B. De Gelder, "Human and animal sounds influence recognition of body language," Brain Res., vol. 1242, pp. 185–190, 2008.

[8] J. Van den Stock, I. Peretz, J. Grezes, and B. de Gelder, "Instrumental music influences recognition of emotional body language," Brain Topogr., vol. 21, no. 3–4, pp. 216–20, May 2009.

[9] H. K. M. Meeren, C. C. R. J. van Heijnsbergen, and B. de Gelder, "Rapid perceptual integration of facial expression and emotional body language," PNAS, vol. 102, no. 45, pp. 16518–23, Nov. 2005.

[10] M. E. Kret, J. J. Stekelenburg, K. Roelofs, and B. de Gelder, "Perception of face and body expressions using electromyography, pupillometry and gaze measures.," Front. Psychol., vol. 4, p. 8, Jan. 2013.

[11] M. Fridin, "Storytelling by a kindergarten social assistive robot: A tool for constructive learning in preschool education," Comput. Educ., vol. 70, pp. 53–64, Jan. 2014.

[12] M. Salem, K. Rohlfing, S. Kopp, and F. Joublin, "A friendly gesture: Investigating the effect of multimodal robot behavior in human-robot interaction," 2011 Ro-Man, pp. 247–252, Jul. 2011.

[13] M. Salem, S. Kopp, I. Wachsmuth, K. Rohlfing, and F. Joublin, "Generation and Evaluation of Communicative Robot Gesture," Int. J. Soc. Robot., vol. 4, no. 2, pp. 201–217, Feb. 2012.

[14] P. Wagner, Z. Malisz, and S. Kopp, "Gesture and speech in interaction: An overview," Speech Commun., vol. 57, pp. 209–232, Feb. 2014.

[15] B. Mutlu, J. Forlizzi, and J. Hodgins, "A Storytelling Robot: Modeling and Evaluation of Human-like Gaze Behavior," 2006 6th IEEE-RAS Int. Conf. Humanoid Robot., pp. 518–523, Dec. 2006.

[16] C. Pelachaud, R. Gelin, J. Martin, and Q. A. Le, "Expressive Gestures Displayed by a Humanoid Robot during a Storytelling Application," pp. 2–5.

[17] Q. Le and C. Pelachaud, "Evaluating an expressive gesture model for a humanoid robot: Experimental results," 2012.

[18] J. Park, H. Song, S. Koo, Y. Kim, and D. Kwon, "Emotions with Modification by Personality," 2010.

[19] J. Xu, J. Broekens, K. Hindriks, and M. A. Neerincx, "Bodily mood expression: Recognize moods from functional behaviors of humanoid robots," in International Conference on Social Robotics, 2013, vol. 8239 LNAI, pp. 511–520.

[20] J. Xu, J. Broekens, K. Hindriks, and M. Neerincx, "Effects of Bodily Mood Expression of a Robotic Teacher on Students," in International Conference on Intelligent Robots and Systems (IROS), 2014, pp. 2614–2620.

[21] J. Montemayor, A. Druin, and J. A. Hendler, "PETS: A personal electronic teller of Stories," Robot. kids Explor. new Technol. Learn., vol. 73, 2000.

[22] N. H. Frijda, The emotion. Cambridge University Press, 1986.

[23] J. A. Russell, "Core affect and the psychological construction of emotion.," Psychol. Rev., vol. 110, no. 1, pp. 145–172, 2003.

[24] C. Beedie, P. Terry, and A. Lane, "Distinctions between emotion and mood," Cogn. Emot., vol. 19, no. 6, pp. 847–878, Sep. 2005.

[25] J. Xu, J. Broekens, K. Hindriks, and M. A. Neerincx, "The relative importance and interrelations between behavior parameters for robots' mood expression," in Proceedings - 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII 2013, 2013, pp. 558–563.

[26] J. Xu, J. Broekens, K. Hindriks, and M. A. Neerincx, "Robot Mood is Contagious: Effects of Robot Body Language in the Imitation Game," in Autonomous Agents and Multi-Agent Systems (AAMAS), 2014, pp. 973–980.

[27] J. Hirth, N. Schmitz, and K. Berns, "Towards social robots: Designing an emotion-based architecture," Int. J. Soc. Robot., no. January, pp. 273–290, 2011.

[28] M. Zecca, Y. Mizoguchi, K. Endo, F. Iida, Y. Kawabata, N. Endo, K. Itoh, and A. Takanishi, "Whole body emotion expressions for KOBIAN humanoid robot — preliminary experiments with different Emotional patterns —," in RO-MAN - IEEE International Symposium on Robot and Human Interactive Communication, 2009, pp. 381–386.

[29] S. Keen, "A theory of narrative empathy," Narrative, 2006.

[30] R. Neumann and F. Strack, "'Mood contagion': the automatic transfer of mood between persons.," J. Pers. Soc. Psychol., vol. 79, no. 2, pp. 211–223, 2000.

[31] J. Tsai, E. Bowring, S. Marsella, W. Wood, and M. Tambe, "A Study of Emotional Contagion with Virtual Characters," in Intelligent Virtual Agents, 2012, pp. 81–88.

[32] C. Creed and R. Beale, "Psychological responses to simulated displays of mismatched emotional expressions," Interact. Comput., vol. 20, no. 2, pp. 225–239, Mar. 2008.

[33] D. C. Berry, L. T. Butler, and F. de Rosis, "Evaluating a realistic agent in an advice-giving task," Int. J. Hum. Comput. Stud., vol. 63, no. 3, pp. 304–327, Sep. 2005.

[34] M. Bradley and P. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," J. Behav. Ther. Exp. Psychiatry, vol. 25, no. 1, pp. 49–59, 1994.

[35] J. Broekens and W. Brinkman, "AffectButton: A method for reliable and valid affective self-report.". International Journal of Human-Computer Studies, 71(6), pp. 641-667, 2013.

[36] E. Aronson, "Back to the Future: Retrospective Review of Leon Festinger's' A Theory of Cognitive Dissonance,'" Am. J. Psychol., vol. 110, no. 1, pp. 127–137, 1997.

978-1-4799-9953-8/15/$31.00 ©2015 IEEE