2017 26th IEEE International Symposium on
Robot and Human Interactive Communication (RO-MAN)
Lisbon, Portugal, Aug 28 - Sept 1, 2017.

Investigating the Real World Impact of Emotion Portrayal through Robot Voice and Motion

Katie Winkle and Paul Bremner[1]

Abstract— **In this paper we investigate robot to human Interpersonal Emotion Transfer (IET) in a real world contextualised human-robot interaction (HRI). IET is an umbrella term which describes the impact of emotions in human-human interaction (HHI). This includes emotion contagion and social appraisal effects. These effects are particularly relevant in domains such as teaching, sports, exercise and healthy eating; domains increasingly targeted by socially assistive robotics. As such, we suggest socially assistive robots may benefit from affective communication in the same way as their human counterparts. We show that emotion recognition from robot voice and motion is possible in explicit validation experiments but does not hold in a socially assistive interaction. Our findings suggest that robot to human IET relies on the human having an expectation for, and hence recognition of, robot emotions; mimicry of valanced motion is not sufficient.**

I. INTRODUCTION

Human beings naturally utilise affective communication in order to effectively communicate their emotional state. The sharing of emotions between persons and the impact this can have on human-human interaction (HHI) is termed Interpersonal Emotion Transfer (IET) [31]. The two main forms of IET are emotion contagion, where an individual 'catches' the emotional state of their interaction partner, and social appraisal, where an individual's judgement is affected by the emotional state of another. For example, listening to a neutral text spoken happily or sadly can induce similar feelings in the listener [30]. Additionally, the same object will be rated differently based on the emotional state of the person presenting it [3].

It has also been demonstrated that IET can have a quantifiable impact on human behaviour (e.g. [2], [42], [28]). In fact, this effect is professionally utilised by those working in assistive human-centred roles which rely on motivation. For example, it has been shown that emotion plays an important part in teaching [27], sports and exercise [40] and healthy eating [7]. The growing area of socially assistive robotics is increasingly concentrated on designing robots for use in such domains. Demonstrated examples include a weight loss coach for the home [18], a memory game playing robot for long term care facilities [24], a robot exercise instructor [13] and a robot language tutor [17]. We suggest that such robots may benefit from affective expression in the same way as their human counterparts.

In this paper we are motivated to investigate the real world impact of affective robot communication using a contextualised HRI experiment. Previous work using a range of robots has demonstrated that, as in human communication, robot emotion expression is possible through voice and motion alone (e.g. [22], [45], [25]), and that robot emotion expression is generally liked by users [37] [24]. However, the quantifiable impact this might have on user behaviour is yet to be investigated. Some initial evidence for robot to human emotion contagion through robot voice and motion has been found [44]; however whether this can be repeated in a socially assistive interaction and its utility in such a context is yet to be investigated.

Here we show that robot voice and motion can portray emotion in within-subject, explicit recognition experiments. However, this does not hold in a between-subject, real world contextualised interaction. We suggest that robot to human IET through voice and motion cannot occur without first achieving emotion recognition; mimicry of valanced motion is not sufficient. Moreover, achieving emotion recognition requires a human expectation that the robot has an emotional state. We also demonstrate that the use of movement can increase the perceived extremity of robot emotions.

II. BACKGROUND AND PREVIOUS WORK

A. Affective Communication in Human-Human Interaction

The phenomenon of IET between humans is still not fully understood; however, it is believed to include social appraisal and emotion contagion effects [31]. There is a general consensus that most emotion contagion is a form of social mimicry, however, it is disputed whether this is based on a physical or empathetic mimicry. The physical mimicry theory states that emotion is generated from mimicking the physical expression, e.g. returning a smile elicits happiness [35]. The empathetic mimicry theory states that physical expression is less important and an individual must perceive another's emotional state and possibly understand its cause in order for emotion contagion to occur [36].

The study of emotion recognition in point light displays generated from dance and acting performances has demonstrated that movement alone can express emotion even if the semantic purpose of that movement is unknown [11], [1], [32]. Laban Movement Analysis (LMA), a multidisciplinary tool for movement analysis, utilises parameters such as weight, space and time to quantify movement for the study of such effects. [20]. Semantic free speech can also portray emotion; for example it has been demonstrated that listening to neutral information spoken in an emotional way, or emotional speech in a foreign language can induce similar emotions in the listener [30], [34]. As with LMA

[1]Both authors are with the Bristol Robotics Laboratory, University of the West of England, Frenchay Campus, Coldharbour Lane, Stoke Gifford, Bristol BS16 1QY k.winkle@bristol.ac.uk and paul.bremner@brl.ac.uk

978-1-5386-3518-6/17/$31.00 ©2017 IEEE

for movement, speech can also be quantitatively described by variation in parameters such as pitch, speed and quality [33] [9]. Models using these parameters can be used to modify neutral speech and motion so that it is conveyed with emotion; this is discussed in the following section.

A range of psychology studies have demonstrated the impact that IET can have on human interaction and behaviour and it is these that provide the motivation for determining whether the same effects might be observed in HRI. It has been demonstrated that emotion expression and emotion contagion can change peoples' rating of the same object [3], help to maintain calm in a dangerous situation [21], increase or decrease trust [12] and improve group task performance [2]. Specific, quantifiable differences in human behaviour have also been demonstrated; for example thirsty individuals pour more or less from a drink jug if exposed to a smiling or frowning face respectively [42] and individuals are more or less likely to accept an offer presented by a smiling or frowning proposer compared to one wearing a neutral expression [28].

B. Emotion Expression through Robot Voice and Motion

Within assistive robotics specifically there are numerous examples of robots that do not have facial expression capabilities (e.g. [29], [18], [14]). Such robots must rely on voice and motion alone in order to portray an emotional state. It is desirable that any emotion expression framework is generic and can be applied to functional robot tasks. This avoids having to implement additional, explicit emotion-specific behaviour. Given the evidence that humans can portray emotion through voice and movement of non-emotional content, this lends itself to use of a voice and movement parameterisation framework.

Examples of such a framework have already been demonstrated [25] [23] [43] and used successfully to generate emotional expressions in a range of robots including the NAO ([23] and [43]). Of these, Lim et al.'s SIRE framework [23] uses the least number of parameters whilst achieving good recognition results and has the benefit of being applicable to both voice and gesture modification. Originally intended for use in telepresence robotics, the four parameters of speed, intensity, regularity and extent, were applied to movement based on a mapping from the same features measured in an actor's speech sample. Implemented on a NAO the resulting motion had an emotion recognition rate of above 60% and, when combined with the original speech sample, lead to improved recognition rates for the emotions of happiness and sadness compared to speech alone.

Lim et al.'s validation of emotion recognition using the SIRE framework did not extend to consider a contextualised interaction. Participants were simply asked to view a range of videos showing the gesturing NAO and to determine its emotional state using a pictorial valence measure. In addition, the impact of the affective movement and any possible contagion effects were not tested for.

C. Related Work

In a rare example of a quantitative study considering robot to human IET, Xu et al. found that participants performed better in a more difficult game when playing with a robot displaying a negative rather than positive mood [44]. The different moods were achieved through a gesture parameterisation framework similar to Lim et al.'s SIRE system [23] but with some additional end-effector specific modifications. The authors used this result to argue that emotion contagion had occurred because of a hypothesised psychological phenomenon that humans undertake certain types of task better when in a negative mood. Aside from task performance itself the impact of this possible contagion and of the affective movements in general was not investigated further.

Other HRI studies have demonstrated that different robot behaviours and communicative cues can impact on human behaviour and perception of the robot (e.g. [15], [8], [38]. Notably, Gockley and Mataric demonstrated that robot behaviour can have an impact on compliance with stroke rehabilitation exercises [14]. However, none of these studies specifically considered affective communication.

III. Research Questions

In this work we investigate the potential consequences of affective robot communication on HRI. We extend previous work by testing for emotion recognition, contagion and impact in a task analogous to a real world application of an assistive robot, i.e. an exercise coach. Specifically, in this context, we ask:

- **RQ1** Can robot emotion be portrayed through voice and motion alone?
- **RQ2** Can robot to human emotion contagion occur?
- **RQ3** What is the impact of affective robot communication?

IV. Materials and Methods

In order to address the research questions a parameterised emotion generation framework was implemented on Softbank Robotics' NAO platform[1]. This framework was validated in a two part emotion recognition study to validate the emotion generation framework (Experiment 1). 16 participants (11 male and 5 female) aged between 17 and 30 (Mean = 25.4 SD = 3.13) were recruited for this using an internal mailing list. The resulting system was utilised in a real world contextualised HRI experiment (Experiment 2) in order to investigate the impact of affective communication. 62 participants (18 male and 44 female) aged between 21 and 60 (Mean = 31.7 SD = 8.72) were recruited for this from the university campus. Participants were offered a five pound Amazon voucher as compensation for taking part in the experiment. The study was approved by the ethics committee of the Faculty of Environment and Technology of the University of the West of England.

[1] https://www.ald.softbankrobotics.com/en/cool-robots/nao

A. Emotional Voice and Motion

A parameterisation framework based on Lim et al.'s Speed, Intensity, Regularity and Extent (SIRE) system [23] was used to make the NAO's movement emotional. The Regularity parameter was not implemented due to to not being utilised in the expression of sadness and happiness which we consider here. Firstly the following key parameters were specified for each movement. In order to achieve natural looking behaviour the movement timing parameters were extracted from an actor filmed playing the role of the robot in the exercise coach interaction:

- $pMin$: a minimum amplitude version of the movement described in the Cartesian co-ordinate set $[x, y, z, \alpha, \beta, \gamma]$
- $pMax$: maximum amplitude version of the movement described in the Cartesian co-ordinate set $[x, y, z, \alpha, \beta, \gamma]$
- $tExt$: time taken to move from base to extended posture based on average from actor video
- $tPos$: time extended posture is held based on average from actor video
- $tRet$: time taken to return from extended to base posture based on average from actor video
- $tMin$: minimum time required to execute extension/return movements for safe operation

Speed & Intensity Speed of the movement is adjusted by modifying the extension and return times passed to the path planner, $tExt$ and $tRet$. These are multiplied by $(1 - S)$ where S is the speed parameter, i.e. as speed is increased the movement time is reduced. A maximum operator is used to compare the adjusted time with the specified minimum time $tMin$ to ensure the movement is not too fast. Intensity is applied by further reducing the extension time, t_1, in the same way as speed, by multiplication with $(1 - I)$ where I is the intensity parameter. This gives the appearance of essentially accelerating extension movements with no change to the return movement, defined by t_2.

$$t_1 = max[(1 - S) * (1 - I) * tExt, tmin] \tag{1}$$

$$t_2 = max[(1 - S) * tRet, tmin] \tag{2}$$

Extent Extent is applied by adjusting the size of the movement between the pre-defined minimum and maximum amplitudes, $pMin$ and $pMax$. A proportion of the difference between them, set by the extent value E, is added to the minimum $pMin$.

$$p = p_{min} + E * (p_{max} - p_{min}) \tag{3}$$

SIE Values Initial SIE values were set based on generic design principles identified by Lim et al. [22] and Xu et al. [43] and then verified through a series of validation experiments as described in the following section. These SIE values are listed in Table I. Figure 1 shows a virtual model of the NAO undertaking the 'Both Up' movement in the sad, neutral and happy conditions, demonstrating the difference in the Extent parameter across conditions.

TABLE I
VALIDATED SIE VALUES

	Sad	Neutral	Happy
Speed	0.1	0.4	0.8
Intensity	0.1	0.4	0.8
Extent	0.1	0.4	0.8

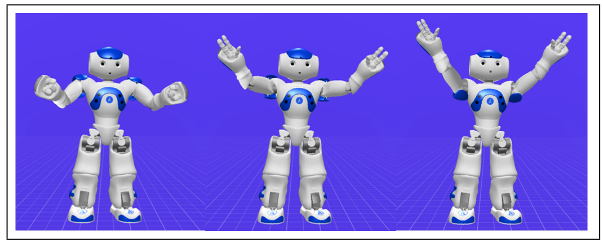

Fig. 1. Sample 'both up' gesture in the sad, neutral and happy conditions demonstrating the different in the Extent parameter across conditions.

Cereproc's CereVoice Engine Text-to-Speech SDK[2] was used to create all speech for the NAO. This text to speech generator was selected because it has built in emotional Synthesis Markup Language (SSML) tags. The 'happy' and 'sad' tags were used to add emotion to pre-determined speech of neutral content.

B. Emotion Expression Validation (Experiment 1)

Two emotion recognition experiments were carried out in order to validate the implemented framework ahead of the main HRI experiment. Stage one of the survey asked participants to watch videos of the moving NAO and identify it's emotional state. Stage two asked participants to listen to voice samples (no visual information was provided) and identify the emotional state of the speaker. Secondly, a new set of 25 participants were recruited to test emotion recognition in the combined voice plus movement condition; representing the final intended behaviour of the NAO. A video demonstrating the voice and motion of different emotional conditions can be accessed online[3]. All gestures and voice clips used in the validation experiments were taken from the main experiment script. Participants were shown 12 samples with 4 each of the neutral, happy and sad condition presented in a different randomised order for each participant. Participants were asked to determine the emotional state of the robot using a 5-point pictorial valence measure.

1) Results: Emotion recognition rates of 60.4%, 68.2% and 65.0% were found in the voice only, movement only and voice + movement validation experiments. These results match those reported in Lim et al.'s original work [23]. Chi-square contingency table analyses were performed to check whether the intended emotion was identified significantly more frequently than chance. This was always found to be the

[2]https://www.cereproc.com/en/products/sdk
[3]https://www.youtube.com/watch?v=9LsWdPf8FH4

case with $\chi^2(1, 192) = 42.8$, $p < 0.001$; $\chi^2(1, 192) = 78.2$, $p < 0.001$ and $\chi^2(1, 300) = 93.9$, $p < 0.001$ respectively. These results validated the emotion expression system and provided the confidence required to use it in the final contextualised HRI experiment. Contingency table analyses were also performed to check the variation in recognition and the number of extreme emotion (i.e. very happy or very sad) choices across the conditions. Overall recognition did not vary significantly but the number of extreme emotion choices, i.e. very happy or very sad, significantly increased between the voice only and voice plus movement conditions $\chi^2(1, 492) = 5.66; p < 0.05$. This is discussed further in Section 6.2.

C. Contextualised HRI Experiment (Experiment 2)

An imitation based robot-led exercise session was designed to provide a contextualised activity for testing the emotional voice and movement framework in a real world human-robot interaction. This was inspired by previous works investigating robot exercise instructors; particularly Fasola and Mataric's work on a robotic arm exercise instructor for the elderly [13]. The choice of gestures (listed in Table 2) was based on Xu et al.'s imitation game [44]. The game element was not implemented due to psychological evidence that winning or losing at a game can influence mood, attitude and likelihood of further game participation [41]. All speech and motion had no emotional content to ensure any resulting IET would be from the robot's voice and motion characteristics only.

The exercise session required participants to undertake rounds of arm exercise in which they had to copy the robot's movements. As physical mimicry is a potential mechanism of IET in humans [35], we aimed to allow for this mechanism to operate with robot IET. The first round of exercise was mandatory, after that participants were asked after each round whether they wished to do another or not. This was designed to give a quantitative measure of how changing the robot's behaviour impacted on participants motivation to exercise. Figure 2 gives a flow chart depiction of the participants' interaction with the robot during the session. The experiment was set up as a between-subject design with participants interacting with the robot in a single emotional condition (happy, sad or neutral) only. Participants were randomly assigned to one of these conditions. In order to avoid a demand effect the experiment was advertised as an investigation into the use of robotic exercise instructors; no mention of affective communication was used on any of the advertising or pre-experiment information material.

1) Experimental Procedure: On entering the experiment room participants were asked to read an experiment information sheet and sign a consent form. A short demo was then given to show the participants what to expect and how to safely interact with the robot. The experimenter then launched the main experiment script and left the room once the robot was seen to be working correctly.

The main exercise session began with the NAO standing up and giving initial instructions followed by the first,

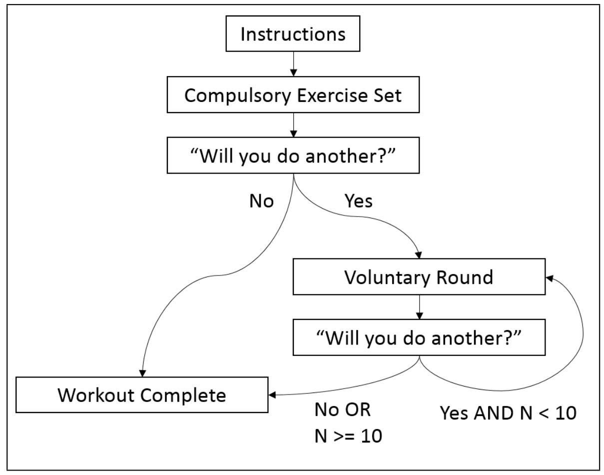

Fig. 2. Simplified interaction flow diagram showing the use of voluntary rounds up to a maximum N = 10

mandatory round of exercise. After this the NAO asked participants if they wished to do another round; if yes then another round was launched, if no the NAO thanked the participant for exercising with it and returned to a crouching position. Each exercise round consisted of 4 movements and an encouraging phrase taken from those listed in Table II.

The mandatory round of exercise was the same for all participants; each voluntary round was chosen randomly in real time from ten alternative, pre-determined options using a random number generator. The gesture and encouragement combinations used in each of the ten options were also generated randomly prior to scripting the robot's control programme. The robot did not monitor the participants' performance in any way. The experimenter returned to the room once the exercise session was complete and the participant was asked to complete an online survey which included all experimental measures.

2) Measures: Emotion recognition of (RQ1) was measured by asking participants to indicate the emotional state of

TABLE II

THE FULL SET OF GESTURES AND ENCOURAGEMENT FROM WHICH EXERCISE ROUNDS WERE GENERATED.

Gesture Instructions	put your right arm up put your right arm down put your left arm up put your left arm down put both arms up put both arms down put your right arm up and left arm down put your left arm up and right arm down
Encouragement	great, well done awesome you're doing really well fantastic excellent nice good job

the NAO during their exercise session. The pictorial valence measure previously used in the validation experiments was used again here. In order to address the question of robot to human emotion contagion (RQ2) two emotional state measures were used; a word completion task and a self assessment. Word completion tasks offer a way to implicitly measure active mental contents through response to an ambiguous task [39]. Such measures have previously been used in HRI to measure death thought accessibility linked to the uncanny valley effect [19] however they have also been used in psychology as an implicit measure of emotional state [10]. The premise of the task is that participants are more likely to complete a valenced word if they are experiencing the matching emotional state, e.g. making J O Y rather than J O G from the word stem J O _ .

The voluntary rounds of exercise completed offers a quantitative measure of the impact of affective robot communication (RQ3). Gockley and Mataric used a similar measure to demonstrate changes in robot behaviour and compliance with an exercise regime [14]. This represents a real world measure of interest when considering the effectiveness of robot exercise instructors. Further, it is known that emotions play a part in the motivation to exercise [40]; therefore if emotion contagion occurred (RQ2) we might expect to see a difference in this measure. Following the word completion task an open comments box offered participants the chance to leave qualitative feedback. Additional qualitative feedback was gathered by informally interviewing participants after the experiment and debriefing was complete.

V. Results and Data Analysis

A. Contextualised HRI Experiment

In contrast with the emotion recognition experiments, participants of the exercise session experiment consistently failed to recognise the intended emotion. Almost all participants judged the robot to be either neutral or happy with roughly an equal split between the two, as shown in Figure 3. The largest difference was in the happy condition, where the robot was judged to be neutral more often than happy, however a chi-square contingency table analysis showed this difference wasn't significant.

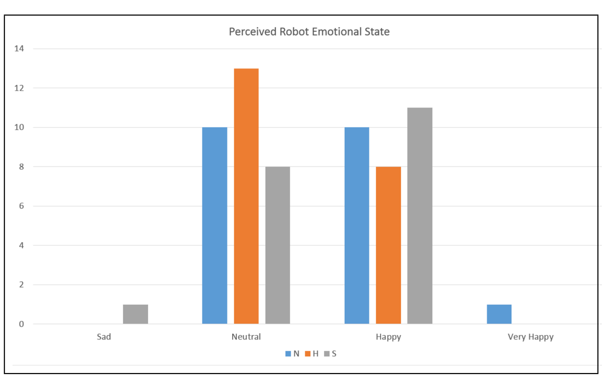

Fig. 3. Robot emotional state as perceived by participants from each emotional condition.

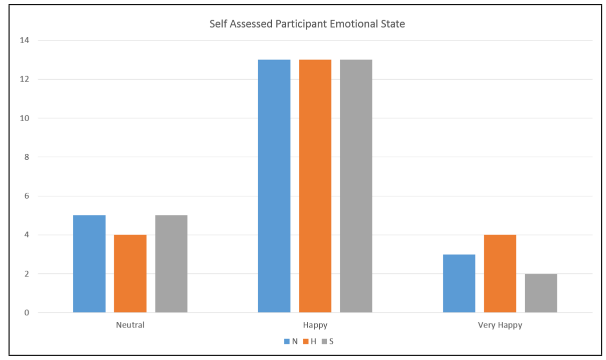

Fig. 4. Participants' self assessed emotional state during the exercise session across conditions.

B. Emotion Contagion

Evidence for emotion contagion was not found in either the word completion or self assessment measures. The number of emotive word completions is given in Table III. Whilst there was an increase between the neutral and emotive conditions, an ANOVA analysis showed this was not significant. Across all conditions the majority of participants indicated they were happy whilst exercising with the robot with a small number saying they were neutral or very happy. This can be seen in Figure 4.

C. Impact of Affective Communication

The mean number of voluntary rounds undertaken by participants in each condition is listed in Table III. This data suggests that people exercised less with the happy robot however a one way ANOVA analysis yielded no significant difference between conditions ($p = 0.14$). One way ANOVA analyses were also carried out for each of the Likert and semantic difference scale survey questions; no significant differences were found in the responses to any question across the three emotion conditions.

D. Qualitative Data

There was very little difference in the qualitative feedback left through the online survey across the emotional conditions. Across all conditions however there was a consistent theme of lack of robot interactivity and responsiveness. Most participants realised the robot was not checking whether they were actually doing the exercises and commented that this reduced the possibility of rapport and the liveliness/lifelikeness of the robot.

Following the experiment participants were asked to read a debrief information sheet explaining the true goal of the study. An informal interview then probed for participants' recognition of the intended emotional state and their thoughts on the use of affective communication in this context. Many in the emotionally valenced conditions were surprised to find out that the robot was supposed to look sad or happy, stating that they weren't expecting to see the robot express emotions and hadn't actually considered whether that was possible.

TABLE III

THE TOTAL NUMBER OF VOLUNTARY ROUNDS OF EXERCISE, EMOTION RECOGNITION AND THE TOTAL NUMBER OF VALENCED WORD COMPLETIONS ACROSS CONDITIONS. POSITIVE AND NEGATIVE WORD COMPLETIONS ARE GIVEN FOR THE NEUTRAL CONDITION DENOTED BY A H AND S RESPECTIVELY.

	Average Rounds	Emotion Recognition	Valenced Word Completions
Neutral	8.67	48%	1.05S / 1.19H
Happy	7.33	38%	1.62
Sad	8.15	5%	1.24

VI. DISCUSSION

A. Emotion Recognition in Robot Gesture & Voice

The most surprising result from this work is the contrast in emotion recognition between the emotion expression validation and final exercise session experiments. Participants of the validation experiments reliably recognised the intended emotion successfully whereas participants undertaking the exercise session failed to do so. Qualitative feedback from the exercise session experiment suggested many participants thought that robots couldn't be emotional and therefore would never have considered that the robot may have an emotional state. This suggests that one key reason for the lack of emotion recognition in the exercise session was participants simply not expecting to see or not being aware that the robot might have an emotional state.

In contrast the validation experiments explicitly asked participants to judge the emotional state of the robot and hence this may have addressed the expectation issue described above. In addition, the validation experiments were within-subject with participants observing multiple examples of each emotional condition. This is consistent with Xu et al. who found, in a very similar within-subject experiment design, that participants were able to distinguish between negative and positive mood in the NAO robot [44]. Undertaking a within-subject design for investigating the real world impact of IET is problematic for three reasons. Firstly, it potentially limits the ecological validity of the study to applications which involve multiple user-robot interactions. Secondly, subjecting participants to multiple emotional states consecutively in a single experiment makes it difficult to measure for emotion contagion across conditions. Finally, it is not clear how subjecting participants to multiple emotional conditions in a single experiment might itself have an impact on participants emotional state.

Considering the contrast between the exercise session, validation experiments and Xu et al.'s results, it would seem that emotion recognition from voice and movement can only be achieved when the participant has some expectation or perception of the robot's emotional expression ability. This may be achieved explicitly through communicating this ability to users (e.g. in the form of explicit questioning) or implicitly through participants observing multiple emotional states. Alternatively this might be achieved by designing an interactive activity which includes an explicit context for the robot's emotional state, or in which the robot's emotional state is more related to the users' task. As discussed in Section 2 we specifically avoided making the robot's emotional state explicit in order to maintain generality and maximise applicability to other, functional socially assistive HRI scenarios. Designing an emotional context that would not itself impact on participants, thus making it impossible to distinguish whether participant emotions were induced by robot to human IET or simply by the interaction context, is non-trivial. This is discussed further in Section 8.

B. Movement & Perceived Emotion Extremity

The finding that combining voice and movement significantly increased the perceived extremity of emotion compared to voice alone is noteworthy. Psychological studies have shown that gesturing is an integral part of HHI [16] [26] and whilst research in the area of robot gesturing has generally shown its utility [6][4] there has been some doubt over gesture speech integration in robot compared to human communicators [5]. The results found in this work however suggest that even functional, emotionally-meaningless movement with valanced movement characteristics can amplify implicit vocal effects. Hence, such modifications to co-speech motion are worth implementing on a robot designed to portray emotion.

C. Contextualised HRI Experiment Results

The lack of emotion contagion found in the exercise session experiment contrasts with the results of Xu et al. who found some evidence for contagion in their gesture imitation game [44]. However, as discussed previously, Xu et al. utilised a within-subject design meaning participants saw a difference in the robot's behaviour between conditions and were therefore likely to have recognised the intended emotional state. This would suggest that emotion recognition is required in order for emotion contagion to occur. This is an interesting result because, based on the HHI psychology discussed in Section 2.1, we might have expected some contagion to occur through participants being asked to physically mimic the robot's valenced movements [35], even if emotion recognition was not achieved. Instead, this result aligns with the empathetic theory of emotion contagion [36] which states that recognition and possibly an understanding of the emotions' cause is required. If this theory holds for robots then the context of this experiment, which gives little reason for the robot to be experiencing an emotion, might be a cause for the lack of contagion. Therefore, whilst it is still unclear in HHI whether contagion is based on mimicry or recognition; this result suggests that, in HRI at least, some level of conscious emotion recognition is required for emotion contagion to occur.

Another possible reason for the lack of contagion is low robot interactivity and responsiveness. In Xu et al.'s imitation game, the robot told the participant whether they were right or wrong after each movement sequence [44]. This work did not utilise any personalised feedback between the robot and participant in order to remove emotional effects related to success or failure as discussed in Section 4.3. As discussed in Section 5.4, this was noted by participants and highlighted as something which reduced the robot's lifelikeness/liveliness, which may have had an impact on the potential for emotion contagion.

The lack of emotion recognition achieved in the exercise session experiment means that conclusions about the impact of affective communication and emotion contagion can not readily be drawn from the results. However, the question of whether affective communication can have an impact on HRI still remains; it may be that if emotion recognition was achieved then the results from the exercise session experiment would have been different. This warrants further investigation and is discussed further in Section 8.

VII. Conclusion

This work investigated affective communication through robot voice and motion and it's potential impact on a real world contextualised HRI scenario. We found that emotion can be successfully conveyed through voice and movement cues on a robot without facial expressions. Additionally, emotions portrayed through voice and movement are perceived as more extreme than voice alone. However, this emotion recognition relies on an expectation of emotional expression, which is not automatic when working with a robot like the NAO. Further, physical mimicry alone is not sufficient to trigger robot to human IET; contrasting with some HHI results [35]. We suggest emotion recognition is required, as according to the theory of empathetic mimicry in HHI [36].

Based on our findings, we recommend two key design principles concerning affective robot communication using voice and motion. Firstly, users must be made aware of the emotional capabilities of the robot. This could be done explicitly, or by demonstrating different emotional states within the same interaction. Secondly, when considering robot design or the implementation of affective communication, it is worthwhile to utilise movement alongside speech in order to express emotional state.

VIII. Limitations and Further Work

This work was limited by the lack of successful emotion recognition in the main HRI experiment, meaning the potential for robot to human IET and the impact this might have on real world HRI still requires further investigation. It would be worthwhile to re-run the exercise session experiment with some modification to address the participant expectation issue such that emotion recognition occurs, to better test robot-human IET. This could be done by exposing participants to multiple emotion conditions in line with Xu et al [44] or by providing some context for the robot's emotional state in line with the HHI literature [36]. However, it may be difficult to create a context which doesn't itself elicit an emotional response from participants. An alternative might be a multiple-interaction scenario in which the emotional expression capability of the robot is demonstrated in one contextualised interaction (e.g. playing a game) but the impact of IET is determined in another.

Acknowledgment

This work was supported by the EPSRC Centre for Doctoral Training in Future Autonomous and Robotic Systems (FARSCOPE) at the Bristol Robotics Laboratory.

References

[1] A. P. Atkinson, W. H. Dittrich, A. J. Gemmell, and A. W. Young. Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, 33(6):717–746, 2004.

[2] S. G. Barsade. The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly*, 47(4):644–675, 2002.

[3] A. P. Bayliss, A. Frischen, M. J. Fenske, and S. P. Tipper. Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition*, 104(3):644–653, 2007.

[4] P. Bremner and U. Leonards. Efficiency of speech and iconic gesture integration for robotic and human communicators-a direct comparison. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1999–2006. IEEE, 2015.

[5] P. Bremner and U. Leonards. Speech and gesture emphasis effects for robotic and human communicators: a direct comparison. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, pages 255–262. ACM, 2015.

[6] P. Bremner, A. Pipe, C. Melhuish, M. Fraser, and S. Subramanian. Conversational gestures in human-robot interaction. In *Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on*, pages 1645–1649. IEEE, 2009.

[7] L. Canetti, E. Bachar, and E. M. Berry. Food and emotion. *Behavioural processes*, 60(2):157–164, 2002.

[8] V. Chidambaram, Y.-H. Chiang, and B. Mutlu. Designing persuasive robots: how robots might persuade people using vocal and nonverbal cues. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, pages 293–300. ACM, 2012.

[9] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor. Emotion recognition in human-computer interaction. *Signal Processing Magazine, IEEE*, 18(1):32–80, 2001.

[10] C. N. DeWall and R. F. Baumeister. From terror to joy automatic tuning to positive affective information following mortality salience. *Psychological Science*, 18(11):984–990, 2007.

[11] W. H. Dittrich, T. Troscianko, S. E. Lea, and D. Morgan. Perception of emotion from dynamic point-light displays represented in dance. *Perception*, 25(6):727–738, 1996.

[12] J. R. Dunn and M. E. Schweitzer. Feeling and believing: the influence of emotion on trust. *Journal of personality and social psychology*, 88(5):736, 2005.

[13] J. Fasola and M. J. Matarić. Robot exercise instructor: a socially assistive robot system to monitor and encourage physical exercise for the elderly. In *RO-MAN, 2010 IEEE*, pages 416–421. IEEE, 2010.

[14] R. Gockley and M. J. Mataric. Encouraging physical therapy compliance with a hands-off mobile robot. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 150–155. ACM, 2006.

[15] J. Goetz and S. Kiesler. Cooperation with a robotic assistant. In *CHI'02 Extended Abstracts on Human Factors in Computing Systems*, pages 578–579. ACM, 2002.

[16] A. Kendon. *Gesture: Visible action as utterance*. Cambridge University Press, 2004.

[17] J. Kennedy, P. Baxter, E. Senft, and T. Belpaeme. Social robot tutoring for child second language learning. In *The Eleventh ACM/IEEE International Conference on Human Robot Interation*, pages 231–238. IEEE Press, 2016.

[18] C. D. Kidd and C. Breazeal. Robots at home: Understanding long-term human-robot interaction. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 3230–3235. IEEE, 2008.

[19] M. Koschate, R. Potter, P. Bremner, and M. Levine. Overcoming the uncanny valley: Displays of emotions reduce the uncanniness of humanlike robots. In *The Eleventh ACM/IEEE International Conference on Human Robot Interation*, pages 359–365. IEEE Press, 2016.

[20] R. Laban. *The Mastery of Movement*. Dance Books Ltd, 2011.

[21] B. Latane and J. M. Darley. Group inhibition of bystander intervention in emergencies. *Journal of personality and social psychology*, 10(3):215, 1968.

[22] A. Lim. *MEI: Multimodal Emotional Intelligence*. PhD thesis, Kyoto University, 2014.

[23] A. Lim, T. Ogata, and H. G. Okuno. Converting emotional voice to motion for robot telepresence. In *Humanoid Robots (Humanoids), 2011 11th IEEE-RAS International Conference on*, pages 472–479. IEEE, 2011.

[24] W.-Y. G. Louie, D. McColl, and G. Nejat. Playing a memory game with a socially assistive robot: A case study at a long-term care facility. In *RO-MAN, 2012 IEEE*, pages 345–350. IEEE, 2012.

[25] M. Masuda and S. Kato. Motion rendering system for emotion expression of human form robots based on laban movement analysis. In *RO-MAN, 2010 IEEE*, pages 324–329. IEEE, 2010.

[26] D. McNeill. *Hand and mind: What gestures reveal about thought*. University of Chicago press, 1992.

[27] D. K. Meyer and J. C. Turner. Discovering emotion in classroom motivation research. *Educational psychologist*, 37(2):107–114, 2002.

[28] P. Mussel, A. S. Göritz, and J. Hewig. The value of a smile: Facial expression affects ultimatum-game responses. *Judgment and Decision Making*, 8(3):381, 2013.

[29] M. Nani, P. Caleb-Solly, S. Dogramadzi, T. Fear, and H. van den Heuvel. Mobiserv: an integrated intelligent home environment for the provision of health, nutrition and mobility services to the elderly. 2010.

[30] R. Neumann and F. Strack. Mood contagion: the automatic transfer of mood between persons. *Journal of personality and social psychology*, 79(2):211, 2000.

[31] B. Parkinson. Interpersonal emotion transfer: Contagion and social appraisal. *Social and Personality Psychology Compass*, 5(7):428–439, 2011.

[32] F. E. Pollick, H. M. Paterson, A. Bruderlin, and A. J. Sanford. Perceiving affect from arm movement. *Cognition*, 82(2):B51–B61, 2001.

[33] K. R. Scherer. Vocal affect expression: a review and a model for future research. *Psychological bulletin*, 99(2):143, 1986.

[34] K. R. Scherer. A cross-cultural investigation of emotion inferences from voice and speech: implications for speech technology. In *INTERSPEECH*, pages 379–382, 2000.

[35] F. Strack, L. L. Martin, and S. Stepper. Inhibiting and facilitating conditions of the human smile: a nonobtrusive test of the facial feedback hypothesis. *Journal of personality and social psychology*, 54(5):768, 1988.

[36] M. Tamietto, L. Castelli, S. Vighetti, P. Perozzo, G. Geminiani, L. Weiskrantz, and B. de Gelder. Unseen facial and bodily expressions trigger fast emotional reactions. *Proceedings of the National Academy of Sciences*, 106(42):17661–17666, 2009.

[37] M. Tielman, M. Neerincx, J.-J. Meyer, and R. Looije. Adaptive emotional expression in robot-child interaction. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, pages 407–414. ACM, 2014.

[38] C. Torrey, S. Fussell, and S. Kiesler. How a robot should give advice. In *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*, pages 275–282. IEEE Press, 2013.

[39] E. Tulving, D. L. Schacter, and H. A. Stark. Priming effects in word-fragment completion are independent of recognition memory. *Journal of experimental psychology: learning, memory, and cognition*, 8(4):336, 1982.

[40] R. J. Vallerand and C. M. Blanchard. The study of emotion in sport and exercise. *Emotions in sport*, pages 3–37, 2000.

[41] J. C. Ward, R. P. Hill, and M. P. Gardner. Promotional games: the effects of participation on mood, attitude, and information processing. *NA-Advances in Consumer Research Volume 15*, 1988.

[42] P. Winkielman, K. C. Berridge, and J. L. Wilbarger. Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Personality and Social Psychology Bulletin*, 31(1):121–135, 2005.

[43] J. Xu, J. Broekens, K. Hindriks, and M. A. Neerincx. Mood expression through parameterized functional behavior of robots. In *RO-MAN, 2013 IEEE*, pages 533–540. IEEE, 2013.

[44] J. Xu, J. Broekens, K. Hindriks, and M. A. Neerincx. Robot mood is contagious: effects of robot body language in the imitation game. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 973–980. International Foundation for Autonomous Agents and Multiagent Systems, 2014.

[45] M. Zecca, N. Endo, S. Momoki, K. Itoh, and A. Takanishi. Design of the humanoid robot kobian-preliminary analysis of facial and whole body emotion expression capabilities. In *Humanoid Robots, 2008. Humanoids 2008. 8th IEEE-RAS International Conference on*, pages 487–492. IEEE, 2008.