Transparent Interactive Reinforcement Learning using Emotional Behaviours*

Georgios Angelopoulos[1[0000−0001−9866−8719]], Alessandra Rossi[1,2[0000−0003−1362−8799]], Gianluca L'Arco[2], and Silvia Rossi[1,2[0000−0002−3379−1756]]

[1] Interdepartmental Center for Advances in Robotic Surgery - ICAROS,
[2] Department of Electrical Engineering and Information Technologies - DIETI,
University of Naples Federico II, Naples, Italy
{georgios.angelopoulos,alessandra.rossi,silvia.rossi}@unina.it

Abstract. This work presents a model for improving transparency during robot learning tasks in Human-Robot Interaction scenarios. Our model puts the human in the learning loop by using two categories of robot's emotional/behavioural reactions, one associated with the learning process of the robot and another elicited as a response to the feedback provided by the user. Preliminary results from a between-subjects study show that people empathized more with a robot expressing its emotions in both the above categories. We noticed a slight increase in the transparency of the robot while it expressed emotions during the learning process and as a response to the user. These findings highlight the importance of emotional behaviours for improving the transparency in the learning systems, which are fundamental for social learning scenarios in future humanoid robotic applications.

Keywords: Interactive Reinforcement Learning · Human-Robot Interaction · Emotional behaviour · Transparency

1 Introduction

The ability of people who lack programming skills (children, older adults, and other non-expert users) to easily teach robots new tasks is becoming critical in domains that involve closer user interactions. As a result, robots need to develop task-related skills with humans as tutors, in similar ways children do, as this will improve the robot's performance and acceptance.

One way for roboticists to provide a robot with learning capabilities is by applying an Interactive Reinforcement Learning (IntRL) algorithm where the human can provide corrections or preferable constraints to enhance the robot's

* This work has been supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 955778, and CHIST-ERA IV COHERENT project "COllaborative HiErarchical Robotic ExplaNaTions", and Italian PON R&I 2014-2020 - REACT-EU (CUP E65F21002920003).

learning [16, 17]. Furthermore, robots are built with anthropomorphic features to allow them to engage people in an interactive learning style that is socially accepted. If a robot's learning behaviour is familiar to people, they will find it more natural to teach it [3].

In this respect, the careful management of robots' behaviour during learning is paramount; when done correctly, people's natural tendencies to anthropomorphize can facilitate and enhance their interaction with robots. However, to better understand how to design the robots' behaviours appropriately, we should understand how human teachers teach their pupils, and how this knowledge can be used to teach robots. In a natural Human-Human environment, a classroom is an emotional place where students frequently express their emotions. For example, students can be excited during studying, hope for success, feel pride in their accomplishments, be surprised at discovering a new solution, or experience anxiety about failing examinations [13]. Another fundamental attribute in educational milieus seems to be teacher's empathy [12]. Arghode et al. [1] showed the significant role of empathy in facilitating the academic development of teachers and students. Baron-Cohen [2] defined empathy as the drive to identify another person's emotions and thoughts. Consequently, just as most adults and children elicit a response to nurture, care, and tutor, robots should elicit a similar response. Indeed, Broekens and Chetouani [4] affirm that the lack of robot transparency has a direct impact on learning. In addition, they highlight the vital link between emotion and expression of the internal state, suggesting that the expression of emotion is a valuable and universal tool, independent of language and species, to transmit one's internal state.

The present work designs, tests, and compares emotional expression mechanisms as a solution for a transparent learning system. In particular, the study explores emotional responses during a robot's learning task based on the progress of the learning and on the certainty of the subsequent actions. We also explore the feedback/reward of humans, and how these behaviours affect their responses. By increasing robots' behaviour transparency, we can design more effective and social robots that are perceived as more acceptable and trustworthy in human-centred environments [14].

2 Related Work

Despite the increasing deployment of humanoid robots in our everyday life, developing transparent interactive learning methods in HRI has just very recently received attention. In particular, robots rarely use emotions to express transparency while learning new tasks from a human teacher.

In a recent study, Hindemith et al. [7] investigated the influence of the feedback type on the user experience of interacting with the different interfaces and the performance of the learning systems. Specifically, they investigated using either absolute scale (e.g., 5-point Likert-scales) or preference-based user feedback (e.g., the participant was shown two movements of the robot and could select which one was better) for an interface to teach a robot a new skill of the game cup-and-ball. While there is no significant difference in the subjective user experience between the conditions, they discovered a significant difference in the learning performance.

Lin et al. [9] proposed an IntRL method to allow a virtual agent to learn from human feedback, such as facial feedback via an ordinary camera and gestural feedback via a leap motion sensor. Their experiments showed that human social signals can effectively improve the learning efficiency of virtual agents. Furthermore, facial feedback recognition error had a larger effect on the agent performance in the beginning training process than in the later training stage.

Suay et al. [16] explored an interactive reinforcement learning approach that enables humans to advise a robot via multiple modalities, such as speech and gestures. Their experimental evaluations in a simulated grid world scenario showed that their method is more robust and converges significantly faster than standard Q-learning algorithms.

Most approaches in the literature generally investigated different feedback types that people use to the virtual learning agent, and they are not focused on how the robot should behave during learning tasks, resulting in a black-box learning system for the users. Therefore, there is still a need for natural and efficient behaviours implemented into humanoid robots during the learning process. Matarese et al. [11] proposed a model to improve the robot's transparency during reinforcement learning tasks by designing non-verbal emotional/behavioural cues into a humanoid robot. Their model considered human feedback as the reward of the RL algorithm, and the robot presented emotional/behavioural responses based on the learning progress. Their results highlighted that people preferred to interact with an expressive robot over a mechanical one. Nevertheless, their model resulted in a misinterpretation when the robot was expressing doubt or uncertainty, and, as a consequence, it negatively affected the robot's transparency. Moreover, the robot's facial expressions were interpreted as a reaction to the user feedback while they were also linked to the learning (certainty/uncertainty) process. Starting from this work, here we present a different model that takes into account also reactions to users' feedback independently of the status of the learning progress and different emotional behaviours. The user perception of every single individual behaviour was previously validated in [15].

3 Methods

During the learning, one of the main challenges is to make the whole process transparent to users, experts or not. This study presents a method where emotions can be used as an effective and transparent solution for communicating the state of the learning process to users. The robot can express emotions that intrinsically represent the current state.

The proposed emotional model relies on the use of four emotions: fear, hope, sadness, and joy. Here, fear and hope are associated with the learning process of the robot, and, therefore, they are elicited during the execution of the robot's actions (e.g., the pointing actions in our application). In this paper, we refer to them as *pointing emotions*. Sadness and joy are, instead, elicited as a response to the feedback provided by the user. In the paper, they are called *feedback emotions*. It is essential to underline that, unlike the approach of Broekens and Chetouani [4], we do not use the pointing emotions as a manifestation of anticipation of a negative or positive adjustment but as the degree of uncertainty in the execution of a specific action a, in a specific state s. Therefore, the pointing emotions represent the robot's degree of certainty about the task execution.

These emotions are expressed based on the CMS model (Color, Motion, Sound) [10]. Following the results of our previous work [15], the robot expresses its emotions through movement and sound, the colour of the LEDs, and also uses its tablet on its chest to make them more recognizable.

3.1 Elicitation of Pointing Emotions

Pointing emotions, E_p, can vary based on intensity. In detail, they fluctuate in a range from maximum fear to maximum hope:

$$E_p = [fear_{high}, fear, fear_{low}, hope_{low}, hope, hope_{high}] \tag{1}$$

Where $fear_{high}$ determines the maximum negative uncertainty, and $hope_{high}$ determines the maximum positive uncertainty. The stimulation of pointing emotions considers the temporal difference error TD (the assessment of how much better or worse a situation just became) and the variation of the temporal difference error Δ_{TD}.

The value of the temporal difference error (initially set at $-\infty$) determines the value of the emotion. When the temporal difference error TD decreases, the valence increases. This behaviour defines the uncertainty of the pointing emotions while executing a specific action a in a state s. The negative uncertainty of an action is mapped as fear, while positive uncertainty as hope. When the TD of a state s converges to 0, the knowledge for the specific state is maximum; in this case, the emotion expressed is $hope_{high}$, and over time, the agent's emotions converge to this emotion.

The intensity of emotions is determined by the variation of the temporal difference error Δ_{TD} (initially set at $-\infty$ as the temporal difference error TD). A significant variation in the temporal difference error Δ_{TD} determines a greater intensity in fear and lower intensity in hope. In contrast, a slight variation in the temporal difference error Δ_{TD} determines a lower intensity in fear and a greater intensity in hope. These emotions were selected based on the work of Tiedens and Linton [18].

In details, let s be a generic non-terminal state, a the action that the agent has chosen to perform; if the absolute value of the relative difference between the new $Q'(s, a)$ and the old $Q(s, a)$, D_q, is less than or equal to 0.1 (so the difference between the two values is at most 10%) to avoid steep changes, then the pointing emotion has a positive value; otherwise negative.

$$valence = \begin{cases} \text{positive, if } D_q = \left| \frac{Q'(s,a)-Q(s,a)}{Q(s,a)} \right| \leq 0.1 \\ \text{negative, otherwise} \end{cases} \tag{2}$$

Table 1: Selection of pointing emotions.

Δ_{TD}	Emotions	
	Positive Valence	**Negative Valence**
< 0.25	$hope_{high}$	$fear_{low}$
$\geq 0.25 \wedge < 0.50$	$hope$	$fear$
≥ 0.50	$hope_{low}$	$fear_{high}$

Once the valence has been established, the absolute value of the relative difference between the new TD' and the old TD, called the Δ_{TD} value, is calculated to determine the specific emotion.

$$\Delta_{TD} = \left| \frac{TD'(s,a) - TD(s,a)}{|TD(s,a)|} \right| \quad (3)$$

Having determined the value and the absolute value of the relative difference between the new and the old TD, Δ_{TD} value, it is possible to identify the corresponding emotion from Table 1.

3.2 Elicitation of Feedback Emotions

Feedback emotions, E_f, just like pointing emotions, can vary based on their intensity; they fluctuate in an interval that goes from maximum sadness to maximum joy:

$$E_f = [sadness_{high}, sadness, sadness_{low}, joy_{low}, joy, joy_{high}] \quad (4)$$

Unlike pointing emotions, feedback emotions do not express information; in fact, they are used only as a direct response to the feedback provided by the user. Furthermore, they are also subject to convergence to an emotion; however, the latter depends on the user's evaluation method. The elicitation of these emotions is based on the feedback provided by the user and on its variation. The feedback value of the user determines the value of the emotion. Negative feedback corresponds to a negative valence, zero feedback corresponds to indifference, while positive feedback corresponds to a positive valence. Once the valence has been established, the specific emotion is determined based on the absolute value of the relative difference between the new R' and old feedback R provided by the user, called Δ_R. In the absence of emotion, the variation of the feedback is indifferent.

$$\Delta_R = \left| \frac{R' - R}{|R|} \right| \quad (5)$$

Having determined the absolute value of Δ_R, it is possible to identify the corresponding emotion from Table 2. Finally, Table 3 presents the adopted CMS model.

Table 2: Selection of feedback emotions

Δ_R	**Emotions**	
	Positive Valence	**Negative Valence**
< 0.25	joy_{low}	$sadness_{low}$
$\geq 0.25 \wedge < 0.50$	joy	$sadness$
≥ 0.50	joy_{high}	$sadness_{high}$

Table 3: Expression of emotions.

Emotions	**Led**	**Movement**	**Sound**
fear low		Indecisive aiming.	Not Applicable.
fear		Indecisive aiming and gaze distortion.	Not Applicable.
fear high		Indecisive pointing.	Not Applicable.
hope low		Fast and decisive aiming.	Not Applicable.
hope		Strong and fast aiming.	Not Applicable.
hope high		Strong and very fast aiming.	Not Applicable.
sadness low		Slightly hunched forward posture.	Discouraged.
sadness		Posture hunched forward.	Discouraged.
sadness high		Very hunched forward posture.	Discouraged.
joy low		Slightly open posture.	Joyful.
joy		Open posture.	Joyful.
joy high		Very open posture.	Joyful.

3.3 The Teaching Scenario

In the proposed study, we use a simple interactive scenario based on the board game Mastermind, invented by Mordecai Meirowitz [5] in which a player, called a decoder, must guess a secret code composed by the opposing player, called the encoder. In our scenario, the robot takes the role of the decoder, and the user is the encoder. The secret code consists of a multi-set of two elements, where each element is a ball of three possible colours. The user's task is limited to choosing the secret code and teaching the robot by evaluating each attempt of the robot to guess the code using feedback $r : r \in [-3, 3] \subset \mathbb{N}$ from a user interface.

The robot's task is to guess the secret code by learning from the ratings provided by the user. Its actions are limited to the choice, and consequent the pointing, of a coloured ball. Based on that, we engineered two behavioural conditions for the robot. Specifically, the conditions are:

- **Condition 1 (C1)**: The robot provides emotional behaviour based on the user reward only (feedback emotions).
- **Condition 2 (C2)**: The robot's behaviour is composed of emotional responses based on the users' rewards (feedback emotion) and emotional behaviour based on the uncertainty of the action (pointing emotion).

We expect that when the emotional expressions of the pointing actions are present in the learning process (C2) will make the robot's behaviours more transparent to the human (**Hypothesis 1**).

In addition, previous studies showed that humans' empathy increases when the robot expresses emotions related to its internal state [8]. Therefore, we expect people to empathize more with the robot in C2 and consequently receive more favourable reward values than in C1 (**Hypothesis 2**).

3.4 Learning Architecture

Considering the interactive scenario described previously, we arrange the coloured balls on the table and enumerate them from 0 to 2; it is possible to formalize the robot's action a as the pointing of the a-th ball. Therefore, the available actions are three (one for each coloured ball available).The state is made up of the set of balls pointed to by the robot, up to that moment, which cannot be greater than the number of balls of the secret code (i.e., 2), and each possible state is a combination of balls with possible repetitions.

The update phase requires a generic sequence of states seq (e.g. $seq = (s_0 = \{\varnothing\}, s_1 = \{1\}, s_2 = \{1, 2\}))$ and a feedback r provided by the user for the sequence of states seq. The Q value of state s is propagated by updating the Q value of all states that can reach s. This set of states is called coverage $C(s)$ (e.g., the coverage of state $\{x, y\}$ is $C(\{x, y\}) = \{\varnothing, \{x\}, \{y\}, \{x, y\}\}$. For every coverage state s_c the $Q(s_c, a)$ values: $s_c \cup \{a\} \in C(s)$ are updated using the classic update equation of the Q-Learning algorithm, with the only particularity that the feedback r is always equal to 0, since the feedback is "assignable" only to the terminal states:

$$Q(s_c, a) = Q(s_c, a) + \alpha \cdot \overbrace{(\gamma \cdot maxQ\left(s_c \cup \{a\}, a'\right) - Q(s_c, a))}^{\text{Temporal Difference error}} \tag{6}$$

The Q value of the reached terminal state s_2 is updated using the following expression:

$$Q(s_2, a') = Q(s_2, a') + \alpha \cdot \overbrace{(r - Q(s_2, a'))}^{\text{Temporal Difference error}} \tag{7}$$

It is important to note that $\forall a : a \in [0, 2] \Rightarrow \gamma \cdot maxQ\left(s_c \cup \{a\}, a'\right) = 0$ since there are no states subsequent to the terminal one. This approach allows the agent to correctly guess the code in a few attempts. Discovering the sequence quickly is essential to avoid participants facing repetitive and potentially dull tests.

4 User Study

A user study was conducted to assess whether the robot's learning process was more transparent to the human teachers when the robot expressed emotions as a reaction to the user feedback or taking into account also the certainty in its action. We designed a between-participant study in a designated environment at the University of Naples Federico II.

4.1 Procedure

Upon arrival, participants were asked to read and sign an informed consent form about the experiment's aims and procedure. Then, the robot and the experimental environment were introduced. Each participant was randomly assigned to one of the conditions. Furthermore, they were told to select a secret code represented by a multi-set of two elements (where each element is a ball of three possible colours) and that they had to evaluate each attempt of the robot to guess the code by providing feedback. Participants were left free to choose their two elements and what feedback to give to the robot. The feedback was provided using the graphical interface by selecting values between -3 and 3.The experimental trial lasted approximately 10-15 minutes.

4.2 Measurement

At the beginning of the study and before the interaction with the robot, participants were asked to complete a questionnaire containing demographic questions (i.e., age, gender, education), their previous experience with robots, and their perception of robots. We also wanted to evaluate the possible negative bias of participants toward robots, so we asked them to answer the following question on a 5-point Likert Scale: "To what extent do you fear that machines will become out of control?".

To understand and measure the individual differences in empathy, we adopted IRI (Interpersonal Reactivity Index) [6], a well-established and validated questionnaire in the social psychology literature, applying it before the experiment. In particular, the participants rated the "Empathic Concern" (EC) subscale, which assesses the feelings of sympathy and concern for unfortunate others, using seven questions on a 5-point Likert Scale.

At the end of the experiment, a questionnaire was administered to the participants ***to measure the transparency of the learning process***. We collected their responses on whether they believed that the robot learned through them ("Do you think the robot learned from your feedback?"), and their expectations ("What was your expectation of the robot after your feedback?"). Finally, we used a 5-point Likert Scale to evaluate to what extent the robot met participants' expectations ("How well does the robot meet your expectations?"). The aforementioned questions evaluate transparency by considering the robot's legibility and predictability attributes.

5 Preliminary Results

We recruited 28 participants (equally distributed in the two conditions) between the University's community, 19 males and 9 females. Their age ranges from 18 to 60 (Mean=28, Std. Deviation=9), and they were not familiar with the setup of the study. The majority of the participants (75%) already had previous experience with robots, while 25% of the participants stated that they had never interacted with robots before. Furthermore, we observed that they had no negative bias towards robots (Max. Value=3, Mean=1.5, Std. Deviation=0.6). For this reason, we did not exclude any participants who successfully participated in the study. Nevertheless, while the limited number of participants in our study makes it difficult to draw definite conclusions, our results, however, indicate some interesting preliminary directions to further investigate.

A Cronbach's α test assessed the internal reliability of the Empathic Concern subscale of the IRI questionnaire, where we found an acceptable value of $\alpha_{EC} = 0.72$. Afterwards, we investigated the mean scores of the EC per each condition, revealing similar mean scores. In particular, in C1, the mean score of Empathic concern was 3.2 ± 0.4, while in C2, we found a mean score of 3.1 ± 0.4.

5.1 System's Transparency

In order to investigate the legibility of the learning system, we analysed participants' responses about their belief that the robot learned from their evaluation. Figure 1 shows that 75% of the respondents in C2 believed that their evaluation helped the learning process of the robot, while there has been a slight decrease of 25% in the number of participants answering positively in C1. In addition, we can also observe a 21% difference between the participants' uncertainty; in C1, half of them stated that they were unsure or replied negatively. However, an Independent Samples T-test did not observe a statistically significant difference, t(22.414) = -1.375, p = 0.18.

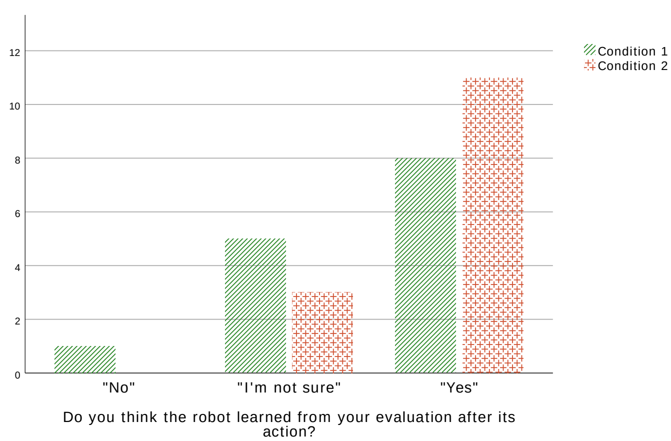

Fig. 1: Humans' confidence about robot's learning per each condition.

Then, we analysed participants' expectations after providing feedback to the robot. Figure 2a shows a slight difference in the participants' answers between the two conditions. Specifically, 64% of the participants in C1 and 78% in C2 believed that it understood what to do in its next move, thanks to their feedback. Nevertheless, we did not observe a statistically significant difference between the two conditions, $t(25.194) = -0.460$, $p = 0.65$

In a secondary exploratory analysis, we examined the participants' responses to the question, "How well the robot met your expectations?". Participants in C1 replied with a mean score of 3.3 and a standard deviation of 0.9, while in C2, we had a mean score of 3.9 with a standard deviation of 0.8. Furthermore, as depicted in Figure 2b the robot in C1 did not fully meet the participants' expectations. In addition, the Independent Samples T-test comparing the two conditions was not statistically significant, $t(25.181) = -1.858$, $p = 0.07$.

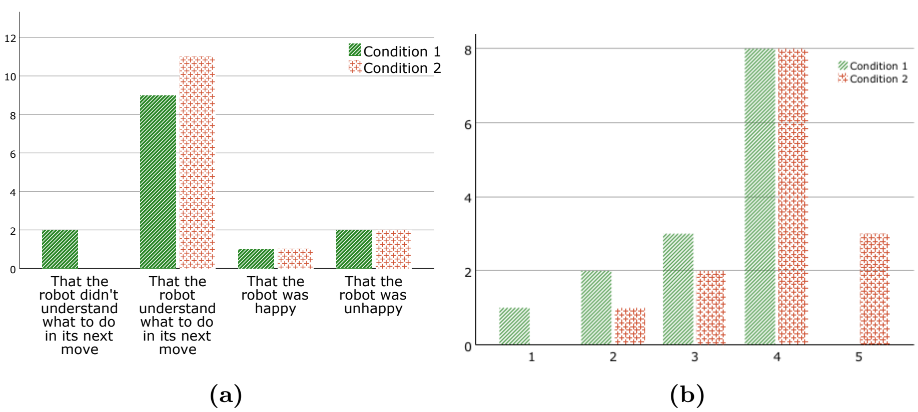

Fig. 2: Participants' expectations. (a) What was your expectation of the robot after your feedback? (b) How well does the robot meet your expectations?

5.2 Participants' Feedback

In terms of the participants' rewards, our study showed that human tutors had a positive bias toward the robot. Our results showed that they opted to reward rather than punish the robot. We can also observe that the robot in C2 received a more favourable reward than in C1 since we have a higher mean value in C2 (Mean= 0.93) than in C1 (Mean = 0.09). Moreover, Figure 3 shows that participants empathized more with the robot in C2 and gave a more favourable reward than in C1. An Independent Samples T-test observed a tendency ($p<0.1$) between the two conditions, $t(22.907) = -1.847$, $p = 0.07$. Therefore, Hypothesis 2 was confirmed. We also noticed that even though we told participants that their feedback could vary between -3 to 3, the participants tended to avoid extreme positive or negative rewards in both conditions. However, participants told the experimenter that they wanted to reward the robot with the most positive feedback (+3) when it learned the sequence. Unfortunately, we did not include a final reward at the end of the learning process. In future works, we would also like to refine our model to consider rewards after learning the sequence.

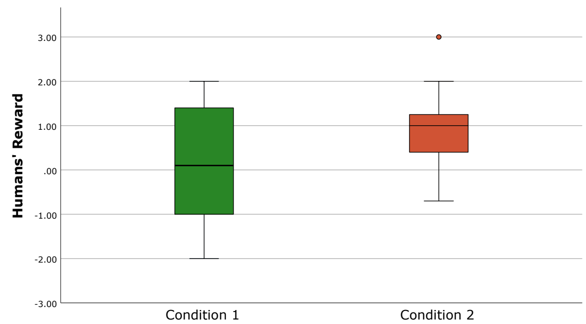

Fig. 3: Humans' reward per each condition.

Observing the participants' behaviours, we noticed that a participant (who was assigned to C2) started with a series of high positive rewards in the first five trials (Mean=2.4) and then continued with a series of smaller rewards (Mean=-0.8). The participant said, "between humans, positive reinforcement works, but not with robots". We assume that this change in participants' strategy is due to the transparency of the learning system. This phenomenon (the alter in participants' approach) occurred in the 14% and 43% of participants in C1 and C2, respectively, (t(23.400) = -1.700, p = 0.1).

6 Conclusions

The work presented in this paper aimed at integrating emotional behaviours into the robot's social learning to improve the transparency of the learning process for human tutors. We compared a robot showing only emotional/behavioural responses based on the user feedback (C1) and a robot expressing emotional/behavioural responses based on the user feedback and emotional expressions based on the certainty of the action (C2). From the experimental results, we observed the transparent effects of the designed human-robot learning system in C2. Furthermore, C2 received more favourable rewards confirming our hypothesis. These findings imply that emotional expressiveness is essential for social robots to interact with people transparently while learning. However, we aim at recruiting a larger and more variate group of participants to confirm the applicability of the phenomenon on a larger scale.

We conclude that our preliminary study offers a starting point for a broader experiment on emotional behaviours during learning with human tutors to achieve transparency and overcome the limitation of previous works. In the future, we will also consider different types of human rewards beyond a user interface that may impact the interaction and the learning process.

References

1. Arghode, V., Yalvac, B., Liew, J.: Teacher empathy and science education: A collective case study. Eurasia Journal of Mathematics, Science and Technology Education **9**(2), 89–99 (2013)
2. Baron-Cohen, S.: The essential difference: The male and female brain. In: Phi Kappa Phi Forum. vol. 85, pp. 23–26 (2005)
3. Breazeal, C.: Designing sociable machines. In: Socially intelligent agents, pp. 149–156. Springer (2002)
4. Broekens, J., Chetouani, M.: Towards transparent robot learning through tdrl-based emotional expressions. IEEE Transactions on Affective Computing **12**(2), 352–362 (2019)
5. Chvátal, V.: Mastermind. Combinatorica **3**(3), 325–329 (1983)
6. Davis, M.H.: Measuring individual differences in empathy: evidence for a multidimensional approach. Journal of personality and social psychology **44**(1), 113 (1983)
7. Hindemith, L., Bruns, O., Noller, A.M., Hemion, N., Schneider, S., Vollmer, A.L.: Interactive robot task learning: Human teaching proficiency with different feedback approaches. IEEE Transactions on Cognitive and Developmental Systems (2022)
8. Kwak, S.S., Kim, Y., Kim, E., Shin, C., Cho, K.: What makes people empathize with an emotional robot?: The impact of agency and physical embodiment on human empathy for a robot. In: 2013 IEEE RO-MAN. pp. 180–185. IEEE (2013)
9. Lin, J., Zhang, Q., Gomez, R., Nakamura, K., He, B., Li, G.: Human social feedback for efficient interactive reinforcement agent learning. In: 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN). pp. 706–712. IEEE (2020)
10. Löffler, D., Schmidt, N., Tscharn, R.: Multimodal expression of artificial emotion in social robots using color, motion and sound. In: 2018 13th ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 334–343. IEEE (2018)
11. Matarese, M., Sciutti, A., Rea, F., Rossi, S.: Toward robots' behavioral transparency of temporal difference reinforcement learning with a human teacher. IEEE Transactions on Human-Machine Systems **51**(6), 578–589 (2021)
12. McAllister, G., Irvine, J.J.: The role of empathy in teaching culturally diverse students: A qualitative study of teachers' beliefs. Journal of teacher education **53**(5), 433–443 (2002)
13. Pekrun, R.: Emotions and learning. Educational practices series **24**(1), 1–31 (2014)
14. Rossi, A., Dautenhahn, K., Lee Koay, K., Walters, M.L.: How social robots influence people's trust in critical situations. In: RO-MAN 2020. pp. 1020–1025 (2020)
15. Rossi, A., Scheunemann, M.M., L'Arco, G., Rossi, S.: Evaluation of a humanoid robot's emotional gestures for transparent interaction. In: International Conference on Social Robotics. pp. 397–407. Springer (2021)
16. Suay, H.B., Chernova, S.: Effect of human guidance and state space size on interactive reinforcement learning. In: 2011 Ro-Man. pp. 1–6. IEEE (2011)
17. Thomaz, A.L., Hoffman, G., Breazeal, C.: Real-time interactive reinforcement learning for robots. In: AAAI 2005 workshop on human comprehensible machine learning. pp. 9–13 (2005)
18. Tiedens, L.Z., Linton, S.: Judgment under emotional certainty and uncertainty: the effects of specific emotions on information processing. Journal of personality and social psychology **81**(6), 973 (2001)