International Journal of Social Robotics (2021) 13:1079–1093
https://doi.org/10.1007/s12369-020-00683-4

A Reinforcement Learning Based Cognitive Empathy Framework for Social Robots

Elahe Bagheri[1] · Oliver Roesler[2] · Hoang-Long Cao[1] · Bram Vanderborght[1]

Accepted: 19 July 2020 / Published online: 20 September 2020

Abstract
Robots that express human's social norms, like empathy, are perceived as more friendly, understanding, and caring. However, appropriate human-like empathic behaviors cannot be defined in advance, instead, they must be learned through daily interaction with humans in different situations. Additionally, to learn and apply the correct behaviors, robots must be able to perceive and understand the affective states of humans. This study presents a framework to enable cognitive empathy in social robots, which uses facial emotion recognition to perceive and understand the affective states of human users. The perceived affective state is then provided to a reinforcement learning model to enable a robot to learn the most appropriate empathic behaviors for different states. The proposed framework has been evaluated through an experiment between 28 individual humans and the humanoid robot Pepper. The results show that by applying empathic behaviors selected by the employed learning model, the robot is able to provide participants comfort and confidence and help them enjoy and feel better.

Keywords Empathy · Reinforcement learning · Personality · Human–robot interaction · Social robot

1 Introduction

The number of robots interacting with humans has increased and some of these robots are interacting with sensitive groups of society, like users with disabilities [8], elders [5,21], people with dementia [9], and children with autism [12], to provide instructions, feedback, and support. Previous studies showed humans prefer robots that show human's social norms, like empathy, which makes robots appear more caring, supportive [7], engageable [23], and friendly [33]. In addition, robots have the opportunity to serve as empathizers in circumstances that humans tend to avoid empathy, e.g., if helping prevents obtaining a desired outcome [11], is exhausting [35] or causes cost [39].

Even though the positive effects of empathy in human-robot interaction (HRI) have been established, a coherent empathic model has not been obtained yet [22], which is mainly due to two reasons. The first reason is that most proposed empathic models consider contextual information and task-related factors, e.g., the user's status in playing a game, i.e., failure or success [25], to conduct empathy, while user's affective state is ignored. In contrast, the model proposed by Churmani et al. [10] considered user's affective state and to determine it asked the users about their feeling through a questionnaire. Whilst the latter approach takes the affective state into account, assuming that users provide accurate information, it interrupts the flow of interaction by asking users to self-report their affective state. To address this problem, the proposed framework has an emotion detection module, which tracks and analyses users' facial expressions to enable a robot to detect users' affective states automatically.

The second reason preventing obtaining a coherent empathy model is defining empathy as a static method that selects empathic behaviors following predefined rules, which cannot be updated during interaction with users [22]. However, empathy is a learnable skill [31] that humans learn throughout their life by interacting with a large number of other humans, and as different humans may desire different types of empathic responses, one's empathic model continuously changes and updates based on the others' reactions. Hence, it is a logical step to let robots also "learn" how to make empathy. This is done in the proposed framework by providing

✉ Elahe Bagheri
elahe.bagheri@vub.be

1 Robotics and Multibody Mechanics Research Group, Vrije Universiteit Brussel and Flanders Make, Brussels, Belgium

2 Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussels, Belgium

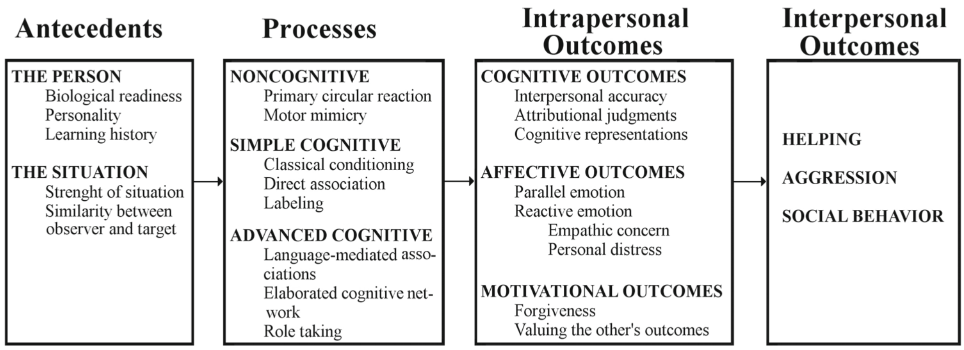

Fig. 1 Cognitive and affective constructs that Davis proposed to explain processes and outcomes of empathy [14]

a learning module, which enables the robot to learn when a certain empathic behavior is appropriate to be applied.

In this study, we try to address the explained problems by developing a coherent empathy model. The main contributions of this work are the use of a facial emotion detection model, which enables the robot to be sensitive and responsive to the users' affective state and the use of a reinforcement learning model to enable the robot to learn the most appropriate empathic behaviors for different situations by trial and error. Applying these two models in the proposed framework leads to a general empathy framework, which is scenario independent.

The remainder of this paper is structured as follows. The applied definition of empathy and its basic constructs are described in Sect. 2. Section 3 presents a review of previous work. The proposed framework is illustrated in Sect. 4. Sections 5 and 6 demonstrate the experimental scenario and obtained results. Finally, Sect. 7 concludes this study.

2 Empathy Definition

Originally empathy has been considered as either a cognitive or an affective phenomenon. Empathy as a cognitive phenomenon is the process where the observer, i.e., empathizer, can *understand* what the other person, i.e., target, is experiencing by taking her perspective but without necessarily experiencing any emotional change. By understanding the target's feelings, the empathizer can provide some reactions more congruence with the target's feeling than her own feeling [38]. In contrast, empathy as an affective phenomenon is an unintentional and uncontrollable process, where the empathizer not only can understand what the target is experiencing but also can *feel* her emotions [14]. Latter, Davis [13] treated empathy as a multidimensional phenomenon that includes both cognitive and affective components as shown in Fig. 1. He defined empathy as a set of constructs that connects the responses of the empathizer to the experiences of the target. These constructs specifically include both the "processes" taking place within the empathizer and the affective and non-affective "outcomes" that result from those processes [14]. The main constructs in his prototype are:

– *Antecedents*, which refer to the attribute of the situation, empathizer or target, e.g., the target's personality;
– *Processes*, which refer to the process by which empathic outcomes are formed;
– *Intrapersonal outcomes*, which refer to cognitive, affective, and motivational empathic outcomes formed in the empathizer that are not necessarily shown to the target;
– *Interpersonal outcomes*, which refer to behavioral empathic outcomes that are shown to the target.

Among different constructs of empathy, developed empathy models in HRI mainly focused on intra- and interpersonal outcomes of empathy. The intrapersonal outcomes of empathy can be *parallel* or *reactive* emotions, thus, previous studies also aimed to develop parallel or reactive empathy models, e.g., [27,33,38].

Parallel outcomes of empathy contain similar emotions that the target is experiencing. In fact, through parallel empathy the empathizer mimics the target's emotions by synchronizing facial expressions, vocalizations, postures, and movements with those of the target [18]. The reactive outcomes of empathy, on the other hand, aim to alter or enhance the target's affective state and can be different from the target's affective state. While parallel outcomes of empathy are more self-oriented, reactive outcomes are focused on the target, thus, reactive outcomes can be considered as a higher level of empathic behavior [27,33].

In this paper we are using David's definition of empathy and since robots are not able to feel as humans can, our focus is on cognitive empathy.

3 Related Work

There are two fundamental approaches for modeling empathy: analytical and empirical. Analytical approaches model empathy by analyzing computational models of empathy and findings of empathy, e.g., findings in the neuroscience of empathy [34]. Empirical approaches, on the other hand, derive empathy directly from observations of "empathy in action" [26]. However, since empathy is a complex concept with no agreement on its exact functionality, there is a lack of computational models of empathy [34]. Therefore, most of the proposed models are derived empirically and attempt to address empathy by designing fixed empathic behaviors for specific scenarios [2]. Such that, robots are trained to evaluate the current affective state of a user in a specific context and then respond to it either by mimicking the observed emotions (parallel empathy) or uttering related predefined sentences (reactive empathy), e.g., [32]. In contrast, Bagheri et al. [4], proposed ACEM, an Autonomous Cognitive Empathy Model, which enables the robot to autonomously decide whether parallel empathy is more adequate or reactive empathy based on the type and intensity of the detected emotion obtained by [3].

Some studies also tried to develop more general models for empathy using learning algorithms. Leite et al. [24] proposed a reinforcement learning based model, where the robot plays chess with users and tries to predict their feelings based on task-related features, e.g., game state and non-verbal behaviors, e.g, gaze direction and smiling, and applies different empathic strategies. The reward function estimates how much a user's positive feeling improved after the robot employed a particular strategy. However, no significant differences were found between the results of the learning based empathic model and static empathic model, which Leite et al. [24] believe is due to the shortness of the interaction, such that users might not have had enough time to realize that adaptation was taking place, and also they believe the time to evaluate the reward function is not defined appropriately.

McQuiggan et al. [27] trained three different Weka [40] models, i.e., Naive Bayes, Decision Tree and Support Vector Machine, in an offline session on biological clues, situational data, and temporal information of empathic interactions of human-controlled virtual agents to predict when and how to exhibit empathic behavior. Afterward, the models were evaluated in an online session, where the virtual empathizer character asks the virtual target character to select one of ten available emotions as her affective state, based on the reported state, the virtual empathizer applies an empathic behavior. The appropriateness of the applied empathic behavior is evaluated by asking the target how she feels after seeing it. Thus, the main disadvantage of this approach is that users need to rate empathic behaviors without feeling real emotions. Additionally, the employed learning models require a training step to know all possible situations that may happen during the experiment and are therefore not able to encounter unseen situations.

Tapus et al. [38] proposed a policy gradient reinforcement learning (PGRL) based model for an assistant robot for a rehabilitation task. The employed model defines some parameters as the robot's overall behavior, which can be optimized by PGRL. The applied reward is the number of exercises performed by the user in the last fifteen seconds, which means that the employed model does not take into account the affective state of the user and can only be used in the scenario described in the study.

Although the described studies tried to learn empathy, their models are restricted to specific scenarios. Therefore, this study proposes a more general reinforcement learning (RL) based empathy framework, which is scenario independent and able to learn which empathic behavior is the most suitable for different emotional states detected in real-time from facial expressions.

4 Proposed Framework

The proposed framework contains three main modules: (1) Emotion Detection module, which detects and recognizes a user's emotion from facial expressions, (2) Reinforcement Learning module, which over time learns to select the empathic behaviors that comfort users in different emotional states, and (3) Empathic Behavior Provider module, which applies selected behaviors to the robot to react to users' emotions. Each individual component is described in more detail in the following subsections.

4.1 Emotion Detection Module

According to Mehrabian [28], 55% of human emotions are conveyed by facial expressions. Hence, we used our facial emotion detection model proposed in our previous work [3] to classify the facial expressions in real-time into six basic emotions defined by Ekman [16], i.e., anger, sadness, surprise, happiness, fear, and disgust. While in proposed scenario (Sect. 5) we did not expect participants to be afraid or disgusted, we only considered the other four emotions. The most dominant emotion during the last ten seconds, i.e., the emotion that was predicted the most, is then given as input to the reinforcement learning module.

4.2 Reinforcement Learning Module

The learning module, illustrated in Fig. 2, is used to enable robots to choose the most appropriate empathic behavior, i.e., the empathic behavior that can comfort the user and alter the user's affective state, in each possible emotional situation. To learn the optimal action-selection policy, the model uses reinforcement learning, more specifically contextual bandit [37]. The number of states is equal to the number of combinations of considered emotions, i.e., four, and considered

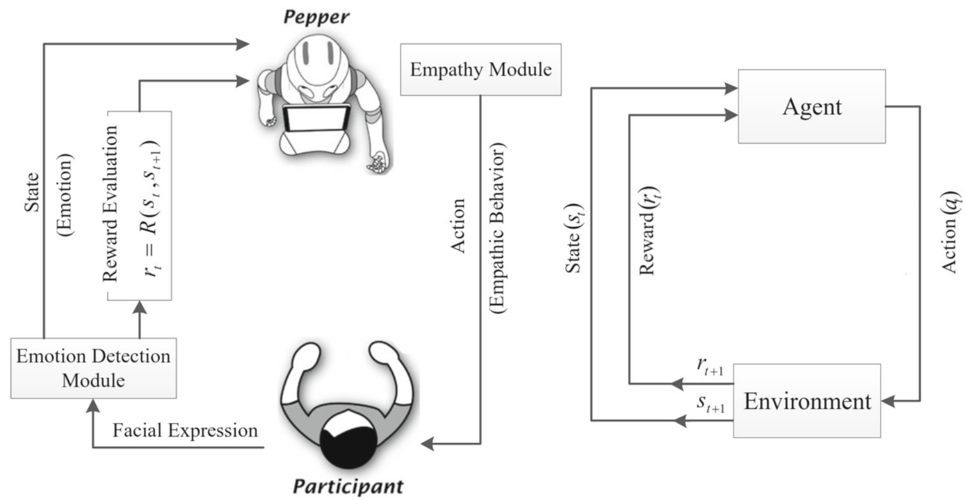

Fig. 2 Architecture of the proposed framework, where the Participant (environment) interacts with the robot Pepper (agent), the Emotion Detection Model recognizes the user's affective state by analyzing Facial Expressions (current state), and Pepper executes an Empathic Behavior (action) selected by the RL model. Afterwards, a reward, which is determined by evaluating the user's new affective state, is given to the applied action

types of personality (Sect. 4.3), i.e., three, leading to a total number of twelve different states.

We considered four categories of empathic utterances as possible actions (Sect. 4.3) so that the robot learns the most appropriate utterance category for each emotional state. Based on the situations the users can encounter in the conducted experiment (Sect. 5), we defined "appropriate responses" as responses that help users to enjoy more. Thus, the robot gets a positive reward when users' emotions change from negative to neutral or positive, or let them stay positive (Table 1).

The Q-table is initialized with zeros. In the beginning, the algorithm starts by selecting random actions, since the Q-values of all utterance types are the same. If the new affective state of the user is undesirable, i.e., the user experiences a negative affective state, the Q-value of the selected action decreases so that the action is less likely to be selected again when the user feels the same affective state in the future. However, if the new affective state is desirable, i.e., the empathic behavior made the user feel better, the Q-value of the selected action increases so that the robot will choose this utterance type with a higher probability, if the user encounters a similar situation again. After performing each action, the Q-table is updated based on the Eq. (1):

Table 1 Illustration of the reward, which depends on the current (S_t) and next (S_{t+1}) emotional states of the user

S_t	S_{t+1} Anger/sadness	Neutral/happiness/surprise
Anger	− 1	1
Sadness	− 1	1
Happiness	− 1	1
Surprise	− 1	1

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r - Q(s,a)], \tag{1}$$

where a is the action taken in state s, r the corresponding reward, and α the learning rate, which determines the importance of prior knowledge. In this study, α is set to 0.1. For exploration ϵ-greedy is used as described by Sutton [37], where ϵ is set to 0.1.

4.3 Empathic Behavior Provider Module

The designed empathic behaviors comprise verbal empathic comments, which are categorized as four sets of utterances (Table 2): *Mimical*, which mimic users' emotions to apply parallel empathy [27], for instance, "I'm happy that you are happy!". *Motivational*, which are a form of reactive empathy that motivate users to pass the current negative emotion [27], such as: "Let it go, look for next round". *Distractional*, which can distract users from negative emotions [6,15], for example: "Do you know what day is today?". *Alleviational*, which try to reduce the user's distress through reactive empathy [20], for instance: "You did your best, don't regret.". While mimical utterances are considered as parallel empathy, motivational, distractional and alleviational utterances are considered as reactive empathy. As shown in Table 2 in each category two utterances are provided for each emotion.

Moreover, personality can affect the characteristics of verbal comments which are preferred by users, e.g., Tapus et al. [38] showed comments which challenge the user are preferred by extrovert users and comments that focus on nurturing praise are preferred by introvert users. In addition, with respect to similarity-attraction principle, extrovert users prefer robots which talk louder, more friendly, and informal, while the introvert users prefer robots which talk lower and more formal [29]. Therefore, we evaluate users' extroversion personality factor through a questionnaire developed by Goldberg [17], where the scores in the extroversion

Table 2 Overview of the used empathic utterances, where for each emotion and category two utterances are considered

Category	Emotion	Utterances	
Mimical	Anger	Is it annoying?	It is annoying me too!
	Sadness	Ooh, it's a pity!	You seem disappointed!
	Happiness	I'm happy you are happy!	I'm happy to play with you!
	Surprise	Seems wonderful!	Amazing, it's fantastic!
Motivational	Anger	Let it go, look for next round.	Hey buddy look forward, next round you will win.
	Sadness	So far you did your best, keep going!	Hey, don't worry, next time you will win.
	Happiness	You seem more beautiful when you smile.	I'm proud of you, you play very well.
	Surprise	Amazing you are very good at this game!	Hey, it's amazing no? Let's go forward.
Distractional	Anger	I preferred if I could play soccer!	I think I'm a lucky robot that can speak with you!
	Sadness	You know I cannot walk when I'm plugged in!	Hey buddy I forget your name! Can you repeat your name?
Alleviational	Anger	I'm sorry, if you are annoyed!	Hey buddy, it's just a short game, take it easy!
	Sadness	Indeed without you, I couldn't pack my bag as this full!	You did your best, don't regret.

For positive emotions, i.e., *Happiness* and *Surprise*, we ignored *Distractional* and *Alleviational* utterances because they only make sense for negative emotions dimension vary between − 5 to 45. Since people with small variations in extroversion dimension have similar personalities, we defined three main categories, i.e., introvert, ambivert, and extrovert with extroversion scores under 16, between 16 and 24, and above 24, respectively. Furthermore, as our platform (Sect. 5) is equipped with some predefined gestures, which allow the robot to talk in a more expressive way, we also applied the most related gesture to each comment.

Once the RL model selects the best utterance category, one of the predefined comments in that category is used randomly. The proposed framework is scenario independent because although the utterances will change based on the context in different scenarios, the defined categories will still be the same, and the goal is finding the category, which leads to better empathic behavior for each situation.

5 Experiment

For this study, 28 participants (19 male and 9 female) were recruited, their mean age was 29 years. Each participant was assigned to one of the defined personality types (introvert = 6, extrovert = 9, and ambivert = 13). All participants were students or university staff and signed an informed consent form before the experiment. At the end of the experiment, participants were compensated with small gifts for their time.

To verify the functionality of the proposed framework, we designed a cooperative interaction scenario, where the participants played a game with the robot while they were not aware of the empathic capability of the robot. The employed robot is Pepper [30], which is a 1.2 meters tall humanoid robot developed by Softbankrobotics. Pepper has twenty degrees of freedom, which allows it to have different postures. It is also equipped with a tablet, microphones, touch sensors, LEDs, and a variety of sensors for multimodal interactions.

The designed game is such that Pepper asks the participant to tell which objects it has put in its magic bag in the correct order. Pepper starts the game by saying: "I put in my bag obj_1", where obj_1 is a randomly selected item from a vocabulary set comprises 42 different objects [1], and asks the participant to help it to remember what it has in its bag. In response, the participant says "You put in your bag obj_1". We used Google Speech API [41] to track the participant's speech and analyze, if the participant had repeated all the objects in the correct order. However, due to sensitivity of the Google Speech API to participants' accents, speech rate, and speech loudness Wizard of Oz confirmed the participant's speech in case Google Speech API fails. Pepper continues the game by adding a new object, and says "I put in my bag, obj_1, obj_2". The game continues until the vocabulary set gets empty or the participant makes a mistake, i.e., the participants knew if they forget one of the objects or say them in the wrong order, they would loose the game. Each session lasted between ten to 45 minutes, depending on how well the participant was in playing the game. We expect users to get happy or surprised when they remember a remarkable list of objects and feel upset or angry when they cannot remember the list of the objects.

Figure 3 shows our experiment setup, in which a user sits in front of Pepper and a camera records the interaction. The emotion detection module tracks participants' facial expressions and sends the detected emotion to the RL model, afterward, Pepper expresses the verbal comment that is selected by the developed RL model with corresponding gesture (Sect. 4.3). Then, the participant's facial expression is tracked again to determine his/her new emotion. Considering this new emotion as the new state, the Q-table is updated based on the reward values defined in Table 1 and Eq. 1.

6 Results and Discussion

To verify the proposed framework, the applied modules are verified individually, i.e., Sect. 6.1 verifies the applied learning module and Sect. 6.2 analyses the users' impression of the robot's empathic behavior provider module through different questionnaires and measurements. Finally, Sect. 6.3 verifies the proposed framework by discussing participants' impressions of the interaction with Pepper.

6.1 Reinforcement Learning Module Verification

Based on the cumulative reward (Fig. 4) over all sessions, i.e., for all participants, the RL model learned the most appropriate empathic behavior for most of the emotional states, i.e., the utterance types that lead the participant to feel more positive. The cumulative reward for individual states (Fig. 5) shows that the most encountered states are when an ambivert or extrovert participant is happy, i.e., 194 or 141 times, respectively. During the experiment negative emotions were not detected so often, e.g, two Sad–Ambivert, seven Sad–Extrovert, four Sad–Introvert, and twelve Anger–Extrovert. Therefore, the employed RL model could not converge to the empathic behaviors that can alter these feelings of the participants, hence, corresponding states are not included in Fig. 5. Table 3 shows the number of occurrence of each defined state.

[1] The vocabulary set contained ordinary items, like pen, apple, tape, spoon, pencil, camera, socks, and scissor.

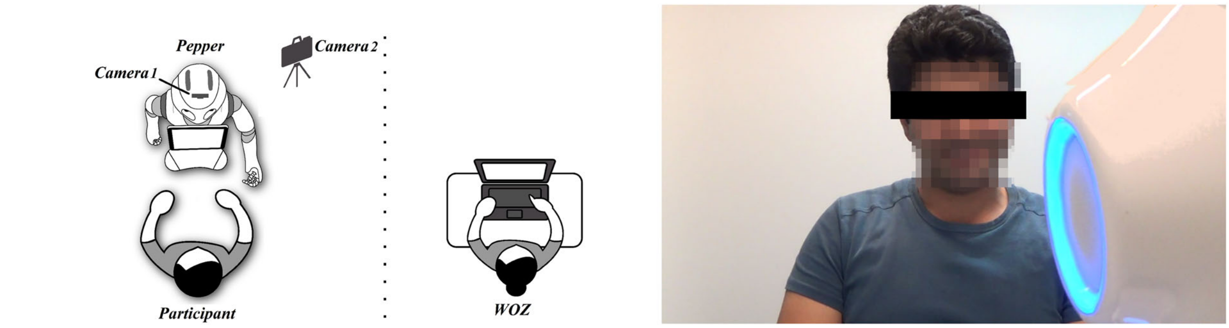

Fig. 3 The setup of the experiment, where the user plays the game with Pepper. The head-mounted camera (Camera 1) is used to record the user's facial expression, while the second camera (Camera 2) is used to record the interaction for offline analysis. The authors obtained the consent for the use of all the photos in this publication

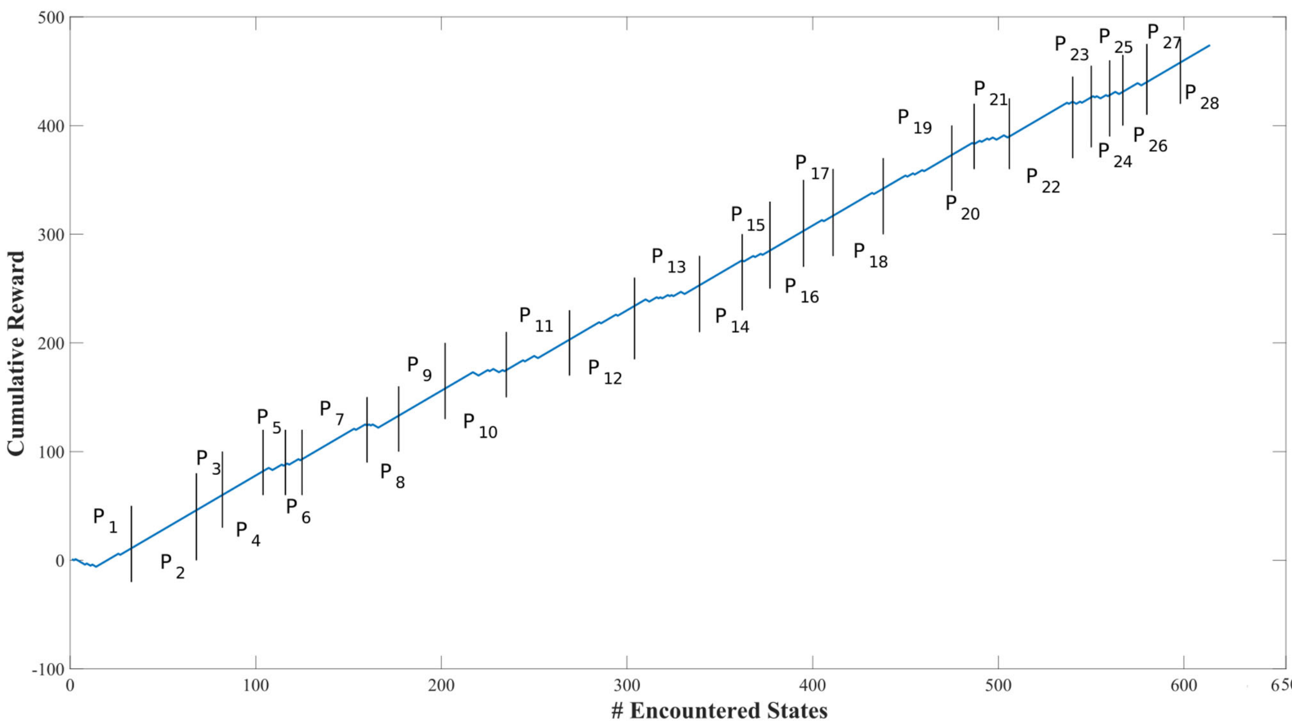

Fig. 4 The cumulative reward that the agent obtained over all participants and states, where P_k, $(k = 1, \ldots, 28)$ shows the session corresponding to the k^{th} participant

According to Fig. 5, for Anger–Ambivert the model required about fifteen situations to learn the empathic utterance type that comfort the participant, i.e., *alleviation*, yet, after a few more situations the participant changed and the learned behavior only sometimes was getting a positive reward for the second participant. After overall forty situations, the participant changed again. In this case, the same behavior as for the first participant received positive feedback from the third participant. For Anger–Introvert the model seemed to converge to *motivational* behavior, but it is not completely clear because this state was encountered only 28 times.

For happiness, the model selected *mimical* empathic behavior at the beginning, which received positive feedback from all participants independent of their personalities. A few cases, where it received negative reward seem to be due to misclassification of the emotion by the emotion detection model or temporary annoyance of participants when the same utterance was used repetitively.

For Surprise–Ambivert and Surprise–Introvert the model converged to *mimical* empathic behavior, while for Surprise–Extrovert it converged to *motivational* behavior. Table 4 summarizes the converged empathic behaviors for all states shown in Fig. 5.

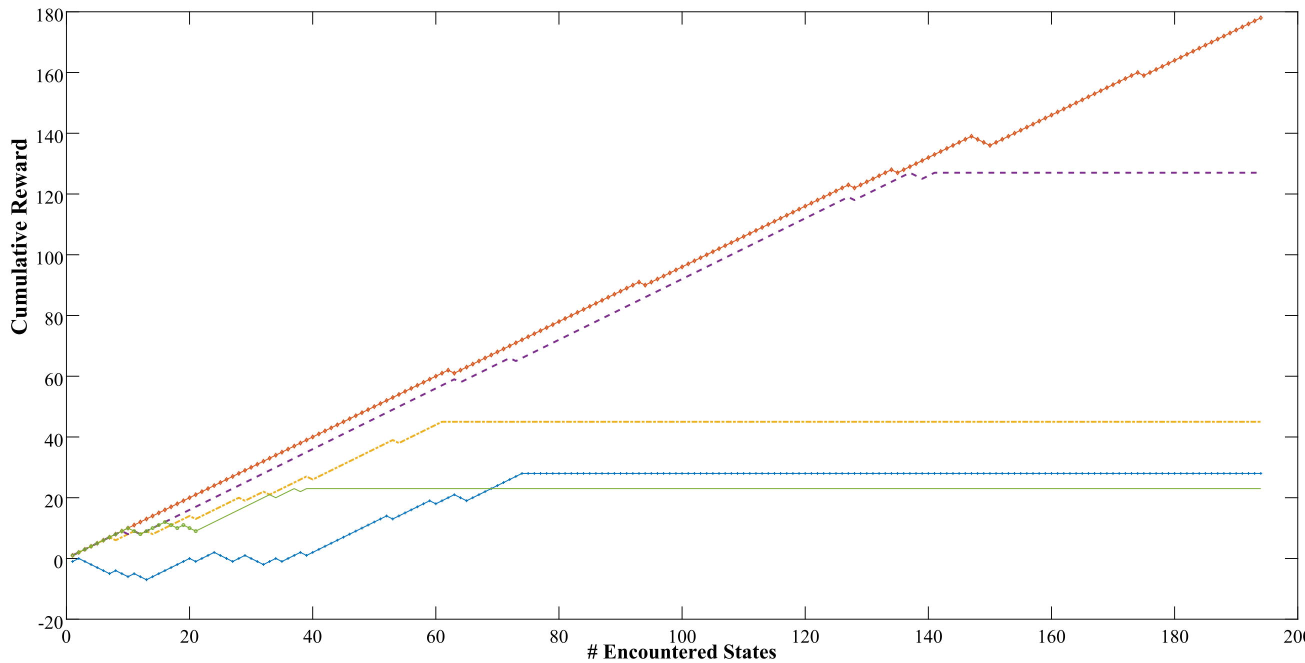

Fig. 5 The cumulative reward that the agent obtained in different states over the number of times it encountered a specific state. The states that the agent encountered for less than fifteen times, i.e., Anger–Extrovert and Sadness for all personalities, are ignored

Table 3 Number of times each emotion was detected during the experiment

Ambivert				Extrovert				Introvert			
Anger	Happiness	Surprise	Sad	Anger	Happiness	Surprise	Sad	Anger	Happiness	Surprise	Sad
74	194	61	2	12	141	39	7	28	45	32	4

Table 4 The type of utterances which the model converged to in different states are indicated by *

Type of action	Ambivert			Extrovert		Introvert		
	Anger	Happiness	Surprise	Happiness	Surprise	Anger	Happiness	Surprise
Mimical		*	*	*			*	*
Motivational					*	*		
Distractional								
Alleviational	*							

For all states, the model sometimes received negative reward after it had already converged to the optimal empathic behavior, which can be a result of one of the following cases: (a) exploration, which can lead to a non-optimal empathic behavior, (b) inaccurate user emotion detection, which can be either that the user's current state is misclassified and an irrelevant empathic behavior is applied or the user's new state is misclassified and the received reward is inaccurate, (c) users annoyance due to repetitive empathic behaviors because users might get annoyed, if they hear the same utterance repeated in a short period of time, and (d) the user changes because there is a probability that different participants prefer different empathic behaviors.

6.2 Empathic Behavior Provider Module Verification

To evaluate the participants' impressions of the proposed empathy framework, we asked them to rate the assertions of three different questionnaires on a scale of one to five with anchors of "strongly disagree" to "strongly agree". To evaluate the internal consistency of the adapted assertions of the applied questionnaires, Mcdonald's ω and Cohen's d are evaluated (Table 6).

The participants' scores to different questionnaires are analysed and shown in terms of average, standard deviation, and 95% confidence intervals. In addition, to evaluate if the obtained results are better than neutral, i.e., being significantly higher than average score (three), we

Table 5 The comparison between the means and standard deviations of our previous study, i.e., ACEM [4] and proposed learning based framework, i.e., RLM

Functionality	Assertions	ACEM		RLM		95% CI
		M	SD	M	SD	
Intimacy	Pepper knew when something was bothering me	**3.45**	1.12	3.32	0.9	[2.55, 3.66]
	Pepper knew when I was upset	**3.4**	1.3	3.28	1.18	[2.27, 3.52]
Emotional Security	If I were worried, Pepper would make me feel better	3.15	1.23	**3.53**	0.79	[2.87, 3.77]
	If I were nervous, Pepper would make me feel calmer	2.97	1.23	**3.43**	0.92	[2.65, 3.64]
	If I were upset, Pepper would make me feel better	3.27	1.2	**3.36**	0.91	[2.65, 3.64]
Social Presence	My thoughts were clear to Pepper	**3.55**	1.22	3.39	0.73	[2.71, 3.65]
	Pepper's thoughts were clear to me	3.57	1.3	**4.36**	0.67	[3.95, 4.69]
	Pepper was influenced by my mood	**4.17**	0.86	3.93	0.81	[3.41, 4.30]
	I was influenced by Pepper's mood	3.02	1.19	**3.78**	0.73	[3.13, 4.08]
Engagement	It was fun watching/playing with Pepper	**3.95**	1.09	3.86	0.88	[3.21, 4.29]
	Pepper made me participate more in the watching/playing	3.27	1.16	**4.5**	0.88	[4.22, 4.77]
	Watching movie/playing game with Pepper caused me real feelings and emotions	2.95	1.26	**3.93**	1.24	[3.34, 4.37]
	I lost track of time while watching/playing with Pepper	3.32	1.27	**3.89**	0.68	[3.41, 4.23]
Perceived Enjoyment (PENJ)	I enjoy the robot talking to me	3.77	1.27	**4.18**	0.93	[3.81, 4.55]
	I enjoy doing things with the robot	3.82	1.07	**4.03**	0.57	[3.68, 4.39]
	I find the robot enjoyable	3.92	1.06	**3.93**	0.93	[3.54, 4.32]
	I find the robot fascinating	3.9	0.83	**4.14**	0.78	[3.68, 4.60]
	I find the robot boring	2.1	1.18	**1.28**	0.46	[1.27, 2.09]
Perceived Sociability (PS)	I consider the robot a pleasant conversational partner	3.22	1.21	**3.25**	0.77	[2.83, 3.67]
	I think the robot is nice	**4.27**	0.8	3.75	0.74	[3.33, 4.17]
	I find the robot pleasant to interact with	**3.67**	0.93	3	0.81	[2.41, 3.59]
	I feel the robot understands me	3.82	1.16	**4.11**	0.97	[3.78, 4.43]

Table 6 The Mcdonald's ω, p value and effect size of the assertions of the friendship, engagement and UTAUT questionnaires, obtained by the proposed RL based framework

Function	Mcdonald		Wilcoxon test		Effect size
	ω	N	V	p	Cohen's d
Intimacy	0.50	2	76.0	.33	0.07
Emotional security	0.53	3	38.5	.10	0.24
Social presence	0.60	4	91.0	< .001	0.91
Engagement	0.61	4	162.5	< .001	0.94
Perceived enjoyment (PENJ)	0.83	5	136.0	< .001	1.13
Perceived sociability (PS)	0.78	4	85.0	.16	0.18

applied Wilcoxon test, since our data suggests a deviation from normality (Table 6). Following we explain each considered questionnaire, its functionalities and obtained results.

6.2.1 Friendship Questionnaire

Applying empathic behaviors in a robot promotes users' positive belief in the robot and its relationship with humans [1,7]. For instance, empathic behaviors can emerge and evolve social connections, such as friendship [25]. Thus, we asked the participants to indicate their level of agreement to nine assertions of friendship questionnaire. The applied assertions selected from [25] are as following:

Intimacy, which refers to "sensitivity to other's needs and states", we used this assertion to see how much our empathic model is sensitive to the users' emotions and needs ($M = 3.3$, $SD = 1.04$ and $p = .33$). The p value is not significant, since the two questions related to intimacy ask, whether the robot is able to understand if the user is upset or worry (Table 5),

Table 7 Categorizing participants based on personality, the number of interactions and the reward group. In the second row the I, A and E refer to the participants' personality type, i.e., Introvert, Ambivert and Extrovert, and in the third row the I and L refer to Initial or Later group, respectively

Participant	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Personality type	A	A	A	E	A	I	E	I	E	E	A	A	E	I	E	A	A	E	A	A	A	A	I	I	E	E	A	I
Initial or later group	I	I	I	I	I	I	I	I	I	I	I	I	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
Number of interactions	33	36	14	23	12	9	45	17	26	34	35	37	37	24	15	18	16	28	29	12	19	36	10	10	17	13	18	16
Reward group	3	1	1	1	3	3	2	3	1	3	3	3	3	1	3	1	2	2	3	1	3	2	3	2	3	3	1	1

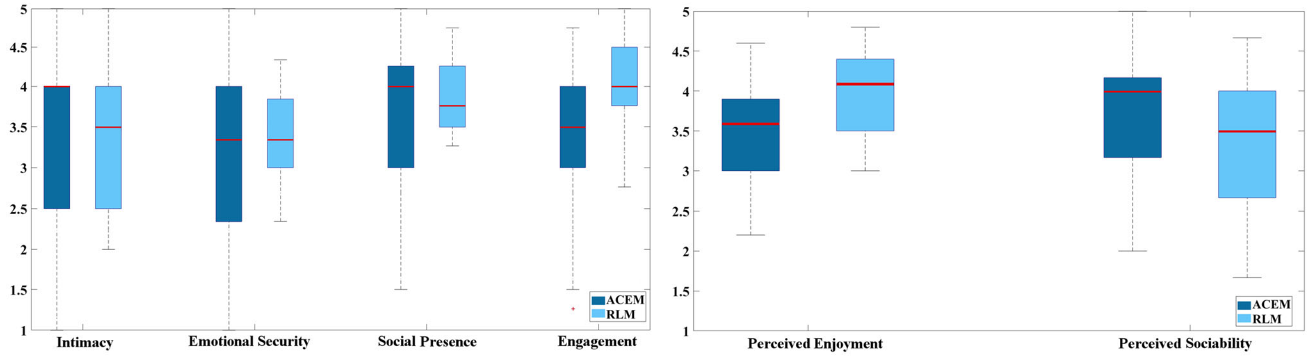

Fig. 6 The comparison between the proposed RL based empathy model, i.e., RLM and our previous study, i.e., ACEM [4]. The red bar in each box plot indicates the median, the box above the median line shows two upper quarters of the data and the box under median line shows two lower quarters of data. The plus marker indicates outlier

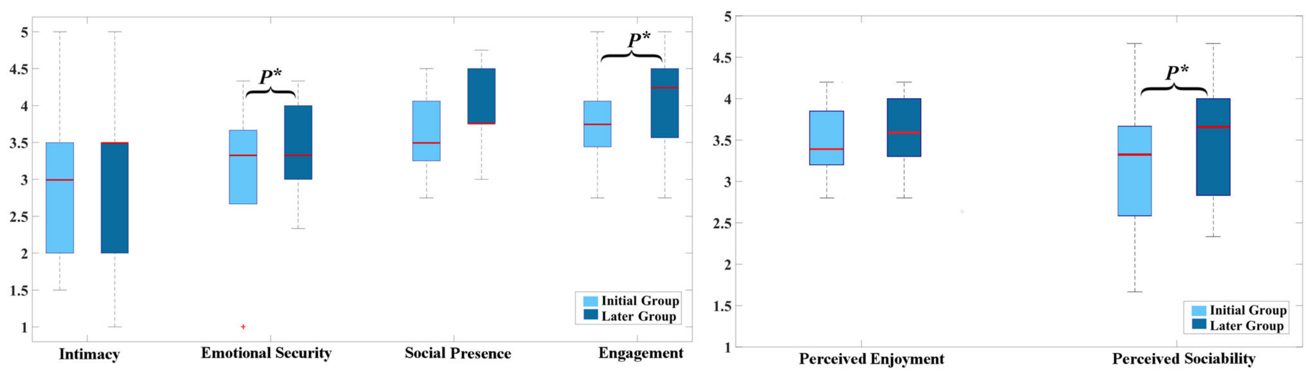

Fig. 7 The comparison of the obtained average score over the applied assertions of the friendship, engagement (left) and the UTAUT (right) questionnaires, for two groups of participants, where the first group interact with the robot, while the learning model was trying to learn and the second group interacted with the robot when the learning module had learned. $p*$ indicates $p < 0.05$, which obtained by Mann–Whitney test

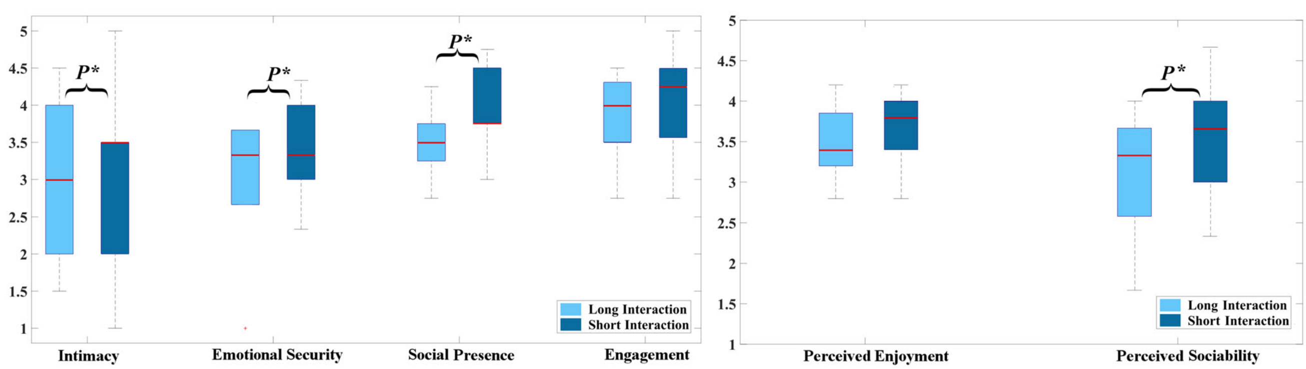

Fig. 8 The comparison of the obtained average score for the applied assertions of the friendship, engagement (left) and the UTAUT (right) questionnaires for two groups of "Long-Interaction" and "Short-Interaction". $p*$ indicates $p < 0.05$, which obtained by Mann–Whitney test

which due to low repetition of sadness and anger in our experiment did not get high scores from participants.

Emotional Security, which refers to "providing comfort and confidence". We use this functionality since we expect the empathic model should be able to comfort participants ($M = 3.44$, $SD = 0.87$ and $p = .1$). The p value is not significant since, similar to Intimacy, the related questions to Emotional Security ask whether the robot is able to make the users feel better or calmer when they are upset or worried

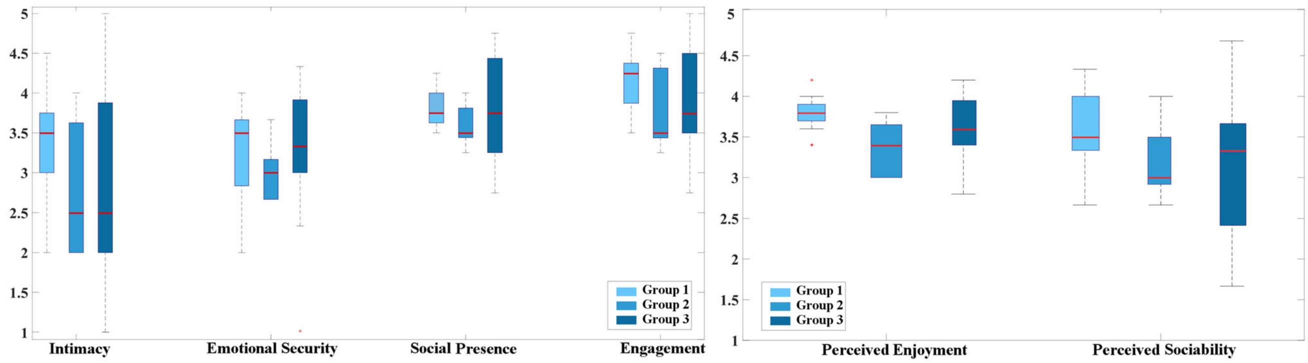

Fig. 9 The comparison of the obtained average score for the applied assertions of the friendship, engagement (left) and the UTAUT (right) questionnaires for three defined groups based on the type of reward

(Table 5), which are the emotions the users did not feel so often during the experiment.

Social Presence, which refers to "the experience of sensing a social entity when interacting with the system", and can show how much the robot's empathic behavior is natural and acceptable, i.e., a higher score in this functionality means the robot is perceived as more humanlike and social. We evaluated this functionality through two sub-functionalities:

Perceived Message Understanding, which refers to "the ability of the user to understand the message from interactant", and a higher score indicates the transparency of the robot's behavior ($M = 3.87$, $SD = 0.7$, and $p < .001$). As the result is significant, one can conclude considered empathic behaviors are appropriate such that participants could understand their relevancy to the situation.

Perceived Affective Interdependence (PAI), which refers to "the extent to which the user's emotional and attitudinal state affects and is affected by the interactant's emotional and attitudinal states". We used this functionality to evaluate the ability of the robot in terms of changing the users' affective state, which is the main goal of empathy ($M = 3.85$, $SD = 0.77$, and $p < .001$). The significance of the obtained results in this functionality indicates that the participants believe the robot gives the impression that it is able to understand their feelings and they got affected by its appropriate reactions. The mean and standard deviation of the participants' scores to the applied friendship assertions are shown in Table 5.

6.2.2 Engagement Functionality

Engagement refers to "the process by which two participants establish, maintain and end their perceived connection" [36] and evaluates the probability of keeping interaction with robots for a longer period of time ($M = 4.04$, $SD = 0.8$ and $p < .001$). Significance of the result shows the robot is able to engage the participants and they are willing to spend more time with the robot. The results are shown in Table 5.

6.2.3 UTAUT Questionnaire

To evaluate the robot acceptance we developed nine assertions of an adapted model of the UTAUT (Unified Theory of Acceptance and Use of Technology) questionnaire [19]. The applied items allowed us to capture the "perceived enjoyment", and "sociability" factors of UTAUT questionnaire. The following applied assertions are selected from [19]:

Perceived Enjoyment, which refers to "Feelings of joy/pleasure associated with the use of the system" [25]. We used this functionality to see how much the users enjoyed interacting with the robot ($M = 3.52$, $SD = 0.75$, and $p < .001$). The obtained results show the participants were enjoying interacting with the robot, which can be seen as the positive effect of provided empathic behavior that users perceived the robot as companionship and can be used to improve the interaction between user and robot.

Perceived Sociability, which refers to "the perceived ability of the system to perform sociable behavior" [25] and enables us to estimate robot's socially ($M = 3.52$, $SD = 0.92$ and $p = .16$). In contrast to the scores of Social Presence that focus on the robot's message and ability to affect users' emotional and attitudinal state, the scores of this functionality are not significant, which can be due to the robot's appearance, i.e., rigid face, robotic voice, and motion (Table 5).

6.3 Framework Verification

A common way to verify proposed models in HRI studies is defining a baseline model and asking participants to rate both models based on different questionnaires. However, defining a baseline for empathic models is challenging. For instance, studies in [24] and [25] defined two types of roles for the robot as "supportive", i.e., proposed empathic model, and "neutral", i.e., baseline model. The supportive model uses utterance like "don't be sad, I believe you can still recover your disadvantage" and the neutral model uses utterance like "bad move", if the user's move in chess play is not good.

Asking participant to rate these two models cannot capture participant's true opinion of the individual models, since one might get a better score just because it is better than the other model and not because it is a good empathy model.

Therefore, in this study we did not define a baseline model that plays a neutral or against role to obtain the individual evaluations of the applied empathy model. Instead, we verified that the proposed model is able to learn the appropriate empathic behaviors by comparing its obtained ratings with the ratings achieved by a model that always applies appropriate empathic behaviors. The latter model, i.e., ACEM, was recently proposed in our previous study [4], which uses the same facial emotion detection module (Sect. 4.1), but a different empathic behavior module, which uses predefined appropriate empathic behaviors. More specifically, ACEM uses reactive empathy, i.e., alleviational and motivational utterances, for negative and strong positive emotions, and parallel empathy, i.e., mimical utterances, for positive and weak negative emotions.

The results in Table 5 and Fig. 6 show that the means and standard deviations are similar for both models indicating that the RL based model has learned the appropriate empathic behaviors for the encountered affective states. For Emotional Security, Perceived Enjoyment, and Engagement, the means are even higher for the RL model, however, this difference is not statistically significant based on an employed Mann–Whitney test. Further investigation are necessary to determine whether learning can lead to more appropriate empathic behaviors than manually predefined ones.

Further, to verify the effect of the learning module on empathic behavior selection, we defined two groups: The first group includes the first half of the participants of each personality type, who interacted with the robot when the applied RL model was not converged and the second group includes the later participants, who interacted with the robot after the model converged (Table 7). Our hypothesis is that the second group scored the robot's empathic behavior higher than the first group since they interacted with the robot when the learning module converged. Figure 7 shows the obtained result, where the second group rated the robot's behavior higher in all considered functionalities. To verify the significance of the obtained results we applied Mann–Whitney test, which shows in three functionalities, i.e., Emotional Security, Engagement and Perceived Sociability, the obtained results by the second group are significantly higher.

Moreover, as shown in Fig. 4, some participants had longer interaction with the robot, while others had shorter interaction (Table 7). To analyse the effect of interaction time on the participants' impression, they are divided into two groups of "Short-Interaction" group, which includes participants that have less than twenty emotional interactions with the robot, i.e., P_3, P_5, P_6, P_8, P_{15}, P_{16}, P_{17}, P_{20}, P_{21}, P_{23}, P_{24}, P_{25}, P_{26}, P_{27}, and P_{28}, and "Long-Interaction" group, which includes the rest of the participants. As shown in Fig. 8, comparing the scores of these two groups "Short-Interaction" group scored the robot's empathic behavior higher. To verify the significance of the obtained results we applied Mann–Whitney test, which shows in four functionalities, i.e., Intimacy, Emotional Security, Social Presence, and Perceived Sociability, the results obtained by the Short-Interaction group are significantly higher. One reason can be that participants in the Long Interaction group heard the applied utterances for many times, which over time can make them bored.

Furthermore, according to Fig. 4, participants can be categorized as three groups: group 1, the participants who always liked the robot's empathic behavior and therefore, the applied empathic behaviors only got positive reward, i.e., P_2, P_3, P_4, P_9, P_{14}, P_{16}, P_{20}, P_{27}, and P_{28}. Group 2, the participants who at the beginning of the interaction liked robot's empathic behavior and robot got positive reward for applying those behaviors but after a while they did not like the robot's empathic behavior any more and robot got negative reward at the end of the interaction, i.e., P_7, P_{17}, P_{18}, P_{22}, and P_{25}, and finally, group 3, who in some cases liked the robot's empathic behavior and in some other cases disliked, i.e., P_1, P_5, P_6, P_8, P_{10}, P_{11}, P_{12}, P_{13}, P_{15}, P_{19}, P_{21}, P_{23}, P_{25}, and P_{26} (Table 7).

We compared the impressions of these three groups based on their scores on the applied questionnaires. Figure 9 shows the group that always liked the robot's behavior, i.e., group 1, has scored the robot's behavior higher in comparison to the two other groups. Interestingly, group 2 scored the robot's behavior less than or equal with group 3. One reason can be that the robot's empathic utterances got repetitive for group 2 after a while, and they did not have the chance to see the robot's behavior adaptation since the interaction stopped shortly afterward. While the participants in group 3, experienced the robot's behavior adaptation, which leads to higher scores by this group. However, Kruskal-Wallis test did not confirm the significance of the obtained result, which might be due to the groups' small and different sample sizes, i.e., nine, five, and fourteen individuals.

7 Conclusion

In this study, we proposed an empathy framework, which determines users' emotional states from facial expressions and uses reinforcement learning to enable robots to select the most appropriate empathic utterance for different affective states. To formalize our task as a RL problem, we considered the combination of four basic emotions and three types of personality as possible states of the environment, defined the type of the provided empathic utterances as possible actions and gave a positive reward, if the user felt more positive after applying the empathic utterance. The results confirmed the ability of the proposed framework in obtaining an empathic policy, which applies the most appropriate type of empathic utterance for each detected state.

To evaluate how much our empathic model was pleasureful for the participants, after the experiment, we asked participants about their impression of interacting with the robot through three different questionnaires of friendship, engagement, and UTAUT. The obtained results from 28 participants showed the effectiveness of the proposed framework in different functionalities, like Social Presence, Perceived Enjoyment, and engagement, i.e., the proposed framework is able to provide participants comfort and confidence and help them enjoy and feel better. In addition, based on the participants' scores in the engagement questionnaire, the proposed model could successfully absorb them.

In future work, we will add speech emotion recognition to the emotion detection module to improve its accuracy for current and next state detection. Additionally, we will investigate the use of Thompson Sampling to handle noisy input data. Finally, we will employ the proposed framework in different scenarios to verify its generalizability.

Acknowledgements The work leading to these results has received funding from Flanders Make Proud (PROgramming by User Demonstration) and the Flemish Government under the program Onderzoeksprogramma Artificiele Intelligentie (AI) Vlaanderen.

References

1. Adams RG, Blieszner R, De Vries B (2000) Definitions of friendship in the third age: age, gender, and study location effects. J Aging Stud 14(1):117–133
2. Asada M (2015) Development of artificial empathy. Neurosci Res 90:41–50
3. Bagheri E, Bagheri A, Esteban PG, Vanderborgth B (2020) A Novel Model for Emotion Detection from Facial Muscles Activity. In: Silva M, Luís Lima J, Reis L, Sanfeliu A, Tardioli D. (eds) Robot 2019: Fourth Iberian Robotics Conference. ROBOT 2019. Advances in Intelligent Systems and Computing, vol 1093. Springer, Cham, pp 237–249. https://doi.org/10.1007/978-3-030-36150-1_20
4. Bagheri E, Esteban PG, Cao HL, De Beir A, Lefeber D, Vanderborght B An autonomous cognitive empathy model responsive to users' facial emotion expressions. ACM Trans Interact Intell Syst (TIIS) (**in press**)
5. Bemelmans R, Gelderblom GJ, Jonker P, De Witte L (2012) Socially assistive robots in elderly care: a systematic review into effects and effectiveness. J Am Med Dir Assoc 13(2):114–120
6. Bentall RP, Haddock G, Slade PD (1994) Cognitive behavior therapy for persistent auditory hallucinations: from theory to therapy. Behav Ther 25(1):51–66
7. Brave S, Nass C, Hutchinson K (2005) Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. Int J Hum Comput Stud 62(2):161–178
8. Brose SW, Weber DJ, Salatin BA, Grindle GG, Wang H, Vazquez JJ, Cooper RA (2010) The role of assistive robotics in the lives of persons with disability. Am J Phys Med Rehabil 89(6):509–521
9. Cao HL, Esteban P, Bartlett M, Baxter PE, Belpaeme T, Billing E, Cai H, Coeckelbergh M, Costescu C, David D (2019) Robot-enhanced therapy: development and validation of a supervised autonomous robotic system for autism spectrum disorders therapy. IEEE Robot Autom Mag 26(2):49–58
10. Churamani N, Barros P, Strahl E, Wermter S (2018) Learning empathy-driven emotion expressions using affective modulations. In: 2018 international joint conference on neural networks (IJCNN). IEEE, pp 1–8
11. Cikara M, Bruneau E, Van Bavel JJ, Saxe R (2014) Their pain gives us pleasure: How intergroup dynamics shape empathic failures and counter-empathic responses. J Exp Soc Psychol 55:110–125
12. Coeckelbergh M, Pop C, Simut R, Peca A, Pintea S, David D, Vanderborght B (2016) A survey of expectations about the role of robots in robot-assisted therapy for children with ASD: ethical acceptability, trust, sociability, appearance, and attachment. Sci Eng Ethics 22(1):47–65
13. Davis MH (1983) Measuring individual differences in empathy: evidence for a multidimensional approach. J Pers Soc Psychol 44(1):113
14. Davis MH (2006) "Empathy" handbook of the sociology of emotions. Springer, New York, pp 443–466
15. Diette GB, Lechtzin N, Haponik E, Devrotes A, Rubin HR (2003) Distraction therapy with nature sights and sounds reduces pain during flexible bronchoscopy: a complementary approach to routine analgesia. Chest 123(3):941–948
16. Ekman P (2002) Facial action coding system (FACS). A human face. https://ci.nii.ac.jp/naid/10025007347/en/
17. Goldberg LR (1992) The development of markers for the big-five factor structure. Psychol Assess 4(1):26
18. Hatfield E, Cacioppo J, Rapson R (1994) Emotional contagion: Cambridge studies in emotion and social interaction. Cambridge University Press, Cambridge. errors-in-variables regression model when the variances of the measurement errors vary between the observations. Statistics in Medicine, vol. 21, pp. 1089–1101
19. Heerink M, Kröse B, Evers V, Wielinga B (2010) Assessing acceptance of assistive social agent technology by older adults: the almere model. Int J Soc Robot 2(4):361–375
20. Hoffman ML (1991) Empathy, social cognition, and moral action. In: Kurtines WL, Gewirtz JL, Lamp JL (eds) Handbook of moral behavior and development. Bd, vol 1, Taylor & Francis, pp 275–301
21. Kachouie R, Sedighadeli S, Khosla R, Chu M-T (2014) Socially assistive robots in elderly care: a mixed-method systematic literature review. Int J Hum Comput Interact 30(5):369–393
22. Legaspi R, Kurihara S, Fukui K.-i, Moriyama K, Numao M (2008) An empathy learning problem for HSI: to be empathic, self-improving and ambient. In: 2008 conference on human system interactions. IEEE, pp 209–214
23. Leite I, Martinho C, Paiva A (2013) Social robots for long-term interaction: a survey. Int J Soc Robot 5(2):291–308
24. Leite I, Pereira A, Castellano G, Mascarenhas S, Martinho C, Paiva A (2011) Modelling empathy in social robotic companions. In: International conference on user modeling, adaptation, and personalization. Springer, pp 135–147
25. Leite I, Pereira A, Mascarenhas S, Martinho C, Prada R, Paiva A (2013) The influence of empathy in human–robot relations. Int J Hum Comput Stud 71(3):250–260
26. McQuiggan SW, Lester JC (2006) Learning empathy: a data-driven framework for modeling empathetic companion agents. In: Proceedings of the fifth international joint conference on autonomous agents and multiagent systems. ACM, pp 961–968
27. McQuiggan SW, Robison JL, Phillips R, Lester JC (2008) Modeling parallel and reactive empathy in virtual agents: an inductive approach. In: Proceedings of the 7th international joint conference on autonomous agents and multiagent systems-volume 1. Interna-

tional foundation for autonomous agents and multiagent systems, pp 167–174
28. Mehrabian A (2017) Nonverbal communication. Routledge, New York
29. Niculescu A, van Dijk B, Nijholt A, Li H, See SL (2013) Making social robots more attractive: the effects of voice pitch, humor and empathy. Int J Soc Robot 5(2):171–191
30. Pandey AK, Gelin R (2018) A mass-produced sociable humanoid robot: pepper: the first machine of its kind. IEEE Robot Autom Mag 25(3):40–48
31. Platt FW, Keller VF (1994) Empathic communication. J Gen Intern Med 9(4):222–226
32. Prendinger H, Ishizuka M (2005) The empathic companion: a character-based interface that addresses users'affective states. Applied Artif Intell 19(3–4):267–285
33. Rasool Z, Masuyama N, Islam MN, Loo CK (2015) Empathic interaction using the computational emotion model. In: IEEE symposium series on computational intelligence. IEEE 2015, pp 109–116
34. Rodrigues SH, Mascarenhas SF, Dias J, Paiva A (2009) "i can feel it too!": Emergent empathic reactions between synthetic characters. In: 2009 3rd international conference on affective computing and intelligent interaction and workshops. IEEE, pp 1–7
35. Shaw LL, Batson CD, Todd RM (1994) Empathy avoidance: forestalling feeling for another in order to escape the motivational consequences. J Pers Soc Psychol 67(5):879
36. Sidner CL, Kidd CD, Lee C, Lesh N (2004) Where to look: a study of human–robot engagement. In: Proceedings of the 9th international conference on intelligent user interfaces. ACM, pp 78–84
37. Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. MIT Press, Cambridge
38. Tapus A, Mataric MJ (2008) Socially assistive robots: the link between personality, empathy, physiological signals, and task performance. In: AAAI spring symposium: emotion, personality, and social behavior, pp 133–140
39. Weisz E, Zaki J (2018) Motivated empathy: a social neuroscience perspective. Curr Opin Psychol 24:67–71
40. Witten IH, Frank E, Hall MA, Pal CJ (2016) Data mining: practical machine learning tools and techniques. Morgan Kaufmann, Burlington
41. Zhang A (2017) Speech recognition (version 3.8). https://github.com/Uberi/speech_recognition/