

Sensors in the Wild: Exploring Electrodermal Activity in Child-Robot Interaction

Iolanda Leite*, Rui Henriques†, Carlos Martinho*, Ana Paiva*

*INESC-ID and IST, Technical University of Lisbon

†DCSE, IST, Technical University of Lisbon

{iolanda.leite,rmch,carlos.martinho}@ist.utl.pt, ana.paiva@inesc-id.pt

Abstract—Recent advances in biosensor technology enabled the appearance of commercial wireless sensors that can measure electrodermal activity (EDA) in user’s everyday settings. In this paper, we investigate the potential benefits of measuring EDA to better understand children-robot interaction in two distinct directions: to characterize and evaluate the interaction, and to dynamically recognize user’s affective states. To do so, we present a study in which 38 children interacted with an iCat robot while wearing a wireless sensor that measured their electrodermal activity. We found that different patterns of electrodermal variation emerge for different supportive behaviours elicited by the robot and for different affective states of the children. The results also yield significant correlations between statistical features extracted from the signal and surveyed parameters regarding how children perceived the interaction and their affective state.

I. INTRODUCTION

One of the main challenges in child-robot interaction is to capture children’s spontaneous and genuine perceptions of the robot and of the interaction in general. Relying solely on traditional survey methods, such as questionnaires or interviews, can often lead to inadequate or misleading evaluations of the interaction because of children’s intrinsic tendency to please adults. This well studied phenomenon in the field of psychology is known as *suggestibility* [1][2], and depends not only on the content and format of the questions, but can also be influenced by other factors such as the age and gender or the interviewer/experimenter [3].

With the recent advances in biosensor technology, small, wireless and non-intrusive sensors are becoming available as commercial products. Examples of such devices include Affectiva’s Q Sensor¹ and edaPlux², which can measure user’s arousal through skin conductance, a form of electrodermal activity (EDA) that increases when users are experiencing excitement, attention, anxiety or high cognitive load, and decreases when experiencing periods of boredom or relaxation [4]. Although research exploring the use of such sensors in HRI is increasing, so far very few studies included children as the target users. Additionally, research shows that children’s electrodermal responses can differ from the average responses measured in adults [5].

Another interesting application for physiological data is real-time affect recognition [4][6]. Here, processed physiological signals are usually combined with other modalities

(e.g., vision, context interpretation, and so on). The HRI community is particularly interested in this application, since socially and affect-aware robots are considered to be more effective in engaging users for longer periods of time [7]. Yet again, most of the existing systems were trained with data collected from adult users [8][9], and children may not respond to certain stimuli as adults do.

The main goal of this paper is to study the potential benefits of measuring EDA to better understand children’s emotional and cognitive processes while interacting with social robots in real-world environments. As outlined above, children’s EDA data can be useful for two distinct purposes. First, physiological signals can be a reliable way of measuring participant’s responses to the robot when compared, for example, to questionnaires or video observation, more likely to be affected by suggestibility. Secondly, EDA data can be used to enhance the real-time prediction of affective states of users, either using solely this type of data or by combining it with other methods.

To achieve this goal, we collected EDA signals from 38 children while interacting with a social robot in their school environment. To analyse the collected data, we propose an expressive method based on temporal windows that enhances local variations by computing angles of electrodermal variation, and considered multiple views of the signal obtained by statistical and geometric metrics such as deviation, average, kurtosis and gradient. Overall, the results suggest that different patterns of electrodermal variation emerge for different supportive behaviours elicited by the robot and for different affective states of the children. The results also yield significant correlations between statistical features extracted from the signal and surveyed parameters regarding how children perceived the interaction and their affective state.

This paper is organised as follows. In the next section, we provide a brief literature review on previous HRI research that employed physiological signals. After that, we present the scenario that was used as case study in this paper. We then present the study methodology and discuss the results. We finalise the paper with some conclusions and implications for future research directions in this area.

II. RELATED WORK

The existing works using physiological signals in HRI are used either to evaluate participant’s responses to the robot or to

¹<http://www.affectiva.com/q-sensor/>

²<http://www.plux.info/EDA>

recognise the user’s affective state. However, the contributions are still not many, especially if considering children as the target users. In fact, to our knowledge, only one work collected children’s physiological information to evaluate the interaction with a social robot [10]. In this work, EDA data was collected from 6 children interacting with an ASIMO robot. Each child had to explain facts about animals to ASIMO, and the robot varied the type of attentive behaviours displayed towards the child depending on the study condition. The results of this study suggest that EDA data can be used to discriminate different types of arousal in children during the interaction.

Most of the previous research on physiological signals in HRI has the main goal of recognising possible affective states of users, since there is strong evidence in the literature that physiological signals can disclose relevant information about the user’s affective states [4][6]. One of such examples is the work of Itoh and colleagues [11], that uses physiological signals to induce stress levels by measuring user’s respiration, heart rate, perspiration, pulse wave and motion. When the robot detects high levels of stress in the user, it generates a different motion in the attempt to reverse this state. In a user study, they verified that the robot’s different motion reduced participants’ stress levels. Similar results were obtained earlier by Rani et al. [8] for the automatic recognition of anxiety, using EDA signals among other physiological signals. By detecting users’ anxiety level, the authors show in another paper that the robot’s responses after recognising user frustration impacted positively the interaction [12]. More recently, Kulić and colleagues [9] used Hidden Markov models to estimate user’s valence and arousal based on heart rate, perspiration and facial muscle contraction signals. The model was trained with data collected during an interaction between 36 participants and a robotic arm.

The use of physiological signals, particularly electrodermal activity, is not only being employed in HRI. It is also becoming increasingly popular in other Human-Computer Interaction fields, such as children’s interaction with virtual agents. For example, Mower et al. [13] reported a study in which they collected autistic children’s EDA signals while interacting with an embodied conversational agent. In the same line of research, Reis and Correia [14] developed The Imaginary Friend, an embodied virtual agent that “walks along” with the user while collecting user’s affective states in a virtual jar. The collected affective states are a combination of electrodermal activity and user’s self-reports in a mobile device.

Despite the growing body of research, there are still many open questions and challenges to address regarding the use of physiological signals in HRI. One of the limitations in the presented studies is that most of them were collected in laboratory environments. In this work, we aim to take the state-of-the-art a step further by presenting the first HRI study collecting children’s EDA signals in a real-world environment.



Fig. 1: Child playing with the iCat wearing the Q Sensor.

III. CASE STUDY

The case study scenario selected to investigate the potential of electrodermal activity in child-robot interaction consists of an iCat robot that acts as an empathic game companion for children (see Fig. 1). The robot is able to play chess with a child using an electronic chessboard³, while employing several supportive behaviours based on the child’s affective state. After every move played by the child, the iCat infers his/her affective state and displays an empathic facial expression in tune with the predicted affective state of the user. Additionally, if the child’s affective state is negative with high probability, the robot also displays socially supportive behaviours in the attempt to reverse this state.

A. Multimodal Affect Recognition

The robot predicts the child’s affective state after every move played by the child. To do so, a multimodal affect detection system that takes into account visual features as well as contextual information of the chess game returns, in real-time, the probability of the user’s positive and negative valence of feeling. The system is based on Support Vector Machines (SVMs) [15].

The *visual* features considered by the affect recognition system (e.g., whether the user is smiling or not, and head direction information) are retrieved by a face tracking software⁴, using a standard webcam placed in front of the child (near the iCat). The *contextual* features (e.g., the state of the game in the child’s perspective, the game evolution since the previous move or if there were any captured pieces in that move) are obtained using the evaluation functions of an open source chess game engine⁵. The affect recognition system was trained with data from Inter-ACT [16], a corpus collected during previous studies of children playing with the iCat robot (for more details on the affect recognition system, please refer to [17]).

³<http://www.dgtprojects.com/site/index.php/products/electronic-boards/usb>

⁴<http://www.seeingmachines.com/product/faceapi/>

⁵<http://www.tckerrigan.com/Chess/TSCP>

Social Support Category	Supportive Behaviours	Examples of implementation in the iCat
Information Support	Suggestion/advice	“Need help? Touch my paw so I can suggest you a move.”
	Teaching	“That was not your best move, because now I can capture your Queen.”
Tangible Assistance	Direct Task	(Play a bad move)
	Tension Reduction	“Shall we start this exercise all over again?”, “I always say, lucky in love, unlucky in chess.”
Esteem Support	Compliment	“That was professionally done!”
	Validation	“Well done, you played what I would have played!”
	Relief of Blame	“Don’t worry, you didn’t have better options.”
	Reassurance	“Something’s not quite right here, but it will get better for sure.”
Emotional Support	Relationship	“I really enjoy playing with you!”
	Understanding	“I understand how you’re feeling, I’ve been through similar situations.”
	Encouragement	“Come on, I still believe in you!”

TABLE I: Examples of the social supportive behaviours implemented in the iCat.

B. Social Supportive Strategies

The supportive behaviours available for the iCat to display were inspired on the framework defined by Cutrona et al. [18], which separates social support in different categories: *information support* (advice or guidance), *tangible assistance* (concrete assistance, for example by providing goods or services), *esteem support* (reinforcing the other’s sense of competence), *emotional support* (expressions of caring or attachment) and *network support* (social integration).

This framework not only provides principles on how to generate different types of supportive behaviours, but also (and more importantly) defines a structure, since similar supportive behaviours are grouped within the same category. For example, there is a clear separation between task-oriented (information support and tangible assistance) and relationship-oriented behaviours (emotional, esteem and network support). Cutrona and colleagues [18] even argue that there are individual differences, such as personality or gender, in the preferences of the support behaviour categories. For this reason, the robot is able to learn, by trial and error, which are the support behaviour categories that are more effective for a particular child, and use behaviours from that category more often when interacting with that child. This adaptation mechanism was implemented using Reinforcement Learning. Since the goal is to maximise the user’s positive valence, the reward function is the difference between the user’s affective state before and after the iCat employed the supportive behaviour [19].

Table I provides examples on how the different support behaviour categories were implemented in the iCat. With the exception of network support (for not being applicable to this particular scenario), the supportive behaviours from all the other remaining categories were adapted to the robot’s behaviour. When the robot selects the support behaviour category more effective for the child, one of the behaviours of that category is selected depending on the context of the game. For example, if the esteem support category is selected, the iCat only congratulates the user for capturing a piece if the user actually captured a piece in the previous move. When several behaviours from the selected category are eligible to display to the user, one of them is selected randomly.

IV. METHOD

We conducted an exploratory study using the case study presented in the previous section. In the study, users interacted with the iCat robot while wearing the Q Sensor⁶, a wireless non-intrusive device for measuring electrodermal activity via skin conductance. The device also contains an accelerometer and measures body temperature. These two metrics were used to detect and correct external interferences of the EDA signal.

A. Participants

The participants who took part in the study were third-grade students from a Portuguese elementary school where children have chess lessons as part of their maths curriculum. Consent forms were distributed in the two classes of the third-grade of the school. We ended up with 38 participants, 15 boys and 23 girls, with ages between 8 and 9 years old.

From the 38 participants, 13 of them (7 girls and 6 boys) had prior exposure to the robot in a long-term study conducted in the same school during five weeks [20]. For these participants, this was the sixth time they were playing with the iCat. We deliberately selected these participants to evaluate whether the EDA signals were different after repeated exposure to the robot. In other words, we were looking for differences in children’s responses after the inherent novelty effect of the first interactions.

B. Protocol

The Q Sensor provides more reliable results after a brief warm up period. Therefore, we considered a warm-up period of approximately 15 minutes, where users were wearing the Q Sensor bracelet before the actual interaction with the robot. During the first ten minutes of this period, they were at their classroom. In the remaining five minutes, they were guided by an experimenter to the room where the experiment was taking place, which was located on the opposite side of the school. This short walk also contributed to an increase in participants’ electrodermal activity.

At the scheduled time, each participant was guided to the room where the experimental setting was installed. After

⁶<http://www.affectiva.com/q-sensor/>

sitting down in front of the iCat, the child was asked to play a chess exercise against the robot and was given a short briefing on how to move the iCat's pieces in the chessboard. The chess exercises consisted in playing a game from a predefined chess position (different from the regular initial setup) suggested by the chess instructor of the school. There were two different exercises, an easy one in which children had advantage in relation to the iCat, and a difficult exercise with a slight advantage to the iCat side. Children were randomly assigned to play only one of these two exercises: 19 played the easy exercise and 19 were assigned to the difficult exercise.

The interaction with the robot lasted, on average, 15 minutes. If after this average time none of the players had checkmated the other, a Wizard-of-Oz option would force the iCat to end the game (either by giving up if it was losing, or by proposing a draw to the child if the robot was in advantage). Apart from this exception, the robot's behaviour was completely autonomous. There was only one experimenter in the room to aid children in case of unexpected events and to end the interaction if needed. After playing with the robot, children were asked to fill a questionnaire in a separate room.

C. Data Collection

The main measure of this study was children's EDA signals, collected during the whole interaction using the Q Sensor. At the beginning of the game we used the event marking feature of the Q Sensor software to synchronise the EDA signals with the remaining data collected during the study.

We also collected interaction logs, generated automatically by the robot during the interaction. These logs contain, among other useful information, the types of supportive strategies employed by the robot in the different moments of the game, and also the probability of the user's affective state given by the multimodal Affect Detection system.

Additionally, users' subjective experience was assessed through a questionnaire. The questionnaire contained several items for each one of the following measures:

- *Engagement*: this measure has been extensively used in HRI and has been defined from several perspectives. We adopted the questionnaire items proposed by Sidner et al. [21], who define engagement as "the process by which two (or more) participants establish, maintain and end their perceived connection".
- *Social Support*: the perception of social support can be defined as "the belief that, if the need arose, at least one person in the individual's circle would be available to serve one or more specific functions" [18]. The questionnaire items measuring perceived support were adapted from the Social Support Questionnaire for Children (SSQC) [22], a self-report questionnaire designed to evaluate children's social support via five different scales: parents, relatives, non-relative adults, siblings, and peers. In this case, we adapted the Peer scale by replacing "a peer" to "iCat".
- *Help and Self-Validation*: these are two dimensions of a Friendship Questionnaire previously employed in a study

where the iCat observes and comments the chess match between two players [23]. With these measures, we intend to evaluate how helpful children perceived the robot, and to what extent they consider the iCat as encouraging and able to help children to maintain a positive image of themselves.

- *Social Presence*: "the degree to which a user feels access to the intelligence, intentions, and sensory impressions of another" [24]. This concept has been widely used to measure people's responses towards different technological artefacts and contains six different sub-scales: *co-presence* (the degree to which the observer believes she/he is not alone), *attentional allocation* (the amount of attention the user allocates to and receives from the robot), *perceived message understanding/perceived affective understanding* (the ability of the user to understand the messages/affective states of the robot), and *perceived affective interdependence/perceived behavioural interdependence* (the extent to which the user's emotional/behavioural states affect and are affected by the robot's emotional/behavioural states).

D. Data Analysis

The collected data per game is composed of four components: *i*) the participant's profile and game conditions, *ii*) the electrodermal activity throughout the game, *iii*) the iCat's interaction logs with the set of performed affective predictions and employed supportive behaviors, and *iv*) the set of surveyed parameters filled by the participant. The collected data was integrated and analyzed according to four major steps.

First, the questionnaire ratings provided by each participant were processed following a simple procedure. For each measure, all the questionnaire items evaluating that measure were grouped and averaged. Since social presence is measured as a combination of five different sub-scales, we used these sub-scales separately. Therefore, we ended up with an average rating, for each user, according to the following measures: *engagement, social support, help, motivation, co-presence, attentional allocation, perceived affective understanding, perceived message understanding, perceived affective interdependence and perceived behavioural interdependence*.

Second, electrodermal signals were analysed with the purpose of understanding to which extent does the signal disclose information related with: *i*) motivational and affective responses, and with *ii*) scores characterizing the overall interaction.

There are two key challenges on mining the EDA signals under continuous interactions as the chess-playing sessions targeted by this work. First, the variations related to the signal behavior are subtle since there is no strong isolated stimulus. Second, depending on the profile of the participant, the signal may present an overall trend to increase (as participants get closer to the end of the game) or to decrease (as the initial expectations and novelty effect fades away). Therefore, local responses to robot-driven actions can be masked by the overall variation.

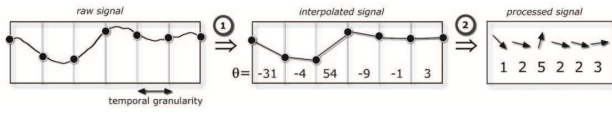


Fig. 2: Illustrative pre-processing of the electrodermal signal using local angles with four symbols

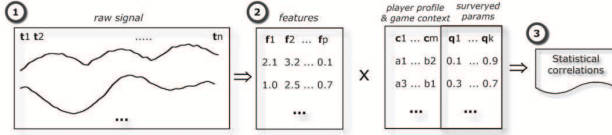


Fig. 3: Feature extraction and analysis from the EDA signal

To surpass these challenges we adopted a novel and expressive pre-processing method to enhance local variations, described in Alg. 1. This method relies on parametrized temporal windows to compute the angles of local electrodermal variation, as illustrated in Fig. 2. The angles are then translated into symbols according to the input number of symbols. Angle cut-off points are defined under a Gaussian assumption.

Input: Float[] eda, Integer nrOfSymbols, Integer deltaT
Output: Integer[] processedEda
 Float[] angles = gaussianCutoffPoints(nrOfSymbols);
 //example: angles ≈ {-36, -12, 0, 12, 36, 90} for nrOfSymbols=6
foreach $i \leftarrow 1$ **to** |eda| **do**
 variation ← eda[i+deltaT]-eda[i];
 slope ← toDegrees(arcTan(variation));
 foreach $j \leftarrow 1$ **to** |angles| **do**
 if slope < angles[j] **then**
 processedEda ← processedEda ∪ j; break;

Algorithm 1: Pre-processing EDA from local variations

Finally, two mining strategies were adopted to analyse the signal. The *first strategy*, illustrated in Fig. 3, retrieves a broad set of features from the discretised and the locally pre-processed signals, and combine them to characterize the overall properties of the interaction. Thus, this strategy considers multiple views over signal’s absolute and local behaviour. The features are mainly statistical and geometric including dispersion metrics such as deviation, distortion and variance; centroid metrics such as simple/harmonic/geometric average; and other key metrics such as Pearson product-moment, kurtosis and gradient.

The *second strategy*, illustrated in Fig. 4, aims to characterize window-based responses related to changes in the participant’s affective states as either a consequence of a changing facial expression or a supportive behaviour of the robot. For this purpose, we used several partitions over the discretised and locally pre-processed signals under two alignment criteria: according to *i*) the affective predictions by the robot, and to *ii*) elicited supportive behaviours by the robot.

Having collected the overall partitions, we conducted a feature analysis (as described for the first strategy) and, additionally, we applied sequential pattern miners to characterize the emerging patterns for each group of partitions. The partitions

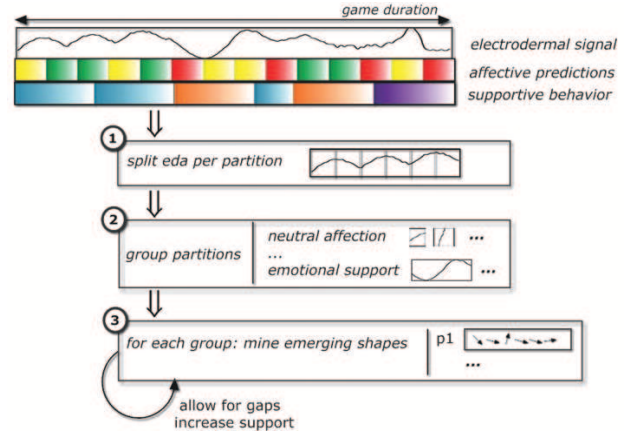


Fig. 4: Emerging patterns from the partitioned EDA signal

(or sequences) were grouped according to the affective prediction (based on the percentage of positive, neutral and negative affect) and to the elicited supportive behaviours (four groups). Alg. 2 describes the adopted algorithm for the analysis of emerging electrodermal sequential patterns across partitions. The adopted sequential pattern miner is uSPAM [25]. The resulting patterns were computed from common precedences under a high level of support that represents a consensus criterion across all signals. Additionally, different patterns were generated under different relaxation conditions (allowing for symbol-gaps). The analysis conducted by the proposed algorithm is able to generalize the common behaviour of the signal under a pattern-based model in relation to the robot’s actions and to changes in the internal affective state of the children.

Input: Signal[] signals, Time[][] affPred, State[][] affType
Output: Pattern[] genPatterns
 Integer affLag ← getAffectiveLag(signals, affectivePredTime);
 Signal[] affSignals;
foreach $i \leftarrow 1$ **to** |signals| /*for each player*/ **do**
 foreach $j \leftarrow 1$ **to** |affPred[i]|-1 /*for each partition*/ **do**
 signal ← signals[i].segment(affPred[i,j], affPred[i,j+1], affLag);
 signal ← localAnglePreProcessing(signal, nrSymbols);
 affSignals[affType[i][j]] ← affSignals[affType[i][j]] ∪ signal;
foreach $i \leftarrow 1$ **to** |affSignals| /*for each affective state*/ **do**
 Bool converge = false;
 Int support = 20%, allowedGaps = 0;
 foreach $j \leftarrow 1$ **to** 10 **do**
 patterns = uspam.run(affSignals[i], support, allowedGaps);
 if converging(patterns) **then** converge = true; break;
 else support += 2%; allowedGaps++;
 if converge **then** genPatterns ← genPatterns ∪ consensus(patterns);

Algorithm 2: Mining generalized EDA patterns

V. RESULTS AND DISCUSSION

The introduced analysis was performed using the collected EDA signals across a chess game (to characterize the experimental setting and surveyed parameters) and over informative partitions of the game (to predict responses to elicited supportive behaviour and the user’s affect).

The use of EDA to disclose information related with experimental conditions is of limited potential. Although EDA features can be used to predict the game result, there were no additional statistically significance for additional conditions, such as the differences regarding the gender or the repeated exposure (first time *versus* repeated users).

A. Questionnaire Measures and Signal Statistical Properties

We ran Pearson’s correlation tests to determine whether the measures obtained in the questionnaires (engagement, help, motivation, social support and the social presence sub-scales) reported in Table II were correlated with any of the overall metrics of the signal (including deviation, distortion, variance, simple/harmonic/geometric average, product-moment, kurtosis and gradient). The most significant results for $N = 38$, considering one-tailed probability, are reported below.

There was a significant relationship between user’s engagement and signal features as kurtosis ($r=.33, p<.05$) and Pearson’s moment-product coefficient between the signal and a linear variable ($r=.29, p<.05$). Since kurtosis measures the “peakedness” of the signal, this result suggests that users who found the interaction with the robot more engaging were the ones who experienced more delineated electrodermal oscillations during the game.

The motivation ratings were significantly correlated with centroids metrics of the signal such as mode ($r=.29, p<.05$), simple mean ($r=.44, p<.01$) and harmonic mean ($r=.74, p<.01$). Since the values of the correlation are positive, this means that higher values of motivation (that is, to what extent children consider the iCat as encouraging and contribute to maintaining a positive image of themselves) are correlated with higher average values of the signal.

Finally, the attentional allocation dimension of social presence was significantly correlated with deviation metrics such as average ($r=-.34, p<.05$), standard ($r=-.35, p<.05$) and squared ($r=-.39, p<.05$) deviations, as well as with variance ($r=-.36, p<.05$). These variables are negatively correlated, which suggests that higher levels of attention are correlated with lower variations in the signal.

The fact that different statistical and geometric metrics of the signal are significantly correlated with different measures suggests that processed EDA data can be used to capture children’s perceptions of the robot at the end of the interaction.

B. Effects of the Robot’s Supportive Behaviours

The partitions framing the participants’ EDA response to the iCat supportive behaviours have the following two properties: assume a lag to exclude the iCat speech (since we are addressing the effect of the behaviour and not evaluating the way iCat approaches the participant), and assume a response duration of 20 seconds. From the over 300 collected partitions, we found significant differences on the emerging patterns for each strategy. The required pre-processing and the mining of the target patterns were, respectively, described in Alg.1 and Alg.2. Under an alphabet of 6-length (being 0 an accentuated negative slope and 5 an accentuated positive slope), we were

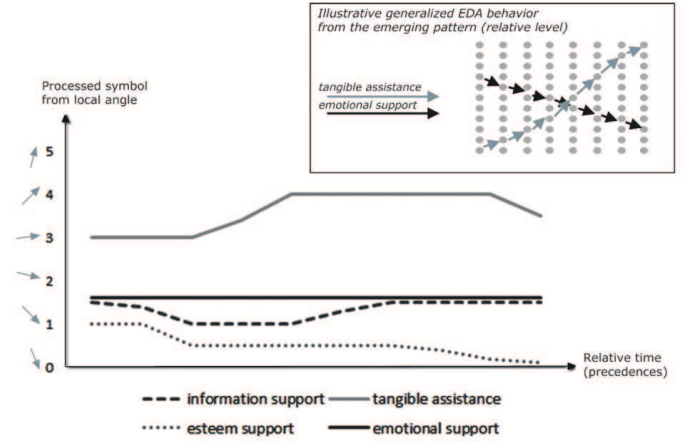


Fig. 5: Emerging sequence patterns for the different supportive behaviours displayed by the robot.

able to generalize the following precedences for each strategy: *i*) information support $\langle(1|2)(1)(1)(1|2)(1|2)(1|2)\rangle$, *ii*) tangible assistance $\langle(3)(3)(3|4|5)(3|4|5)(3|4|5)(3|4)\rangle$, *iii*) esteem support $\langle(1)(0|1)(0|1)(0|1)(0|1)(0)\rangle$, and *iv*) emotional support $\langle(1|2)(1|2)(1|2)(1|2)(1|2)\rangle$. These patterns are graphically illustrated in Fig 5.

This generalized behaviour is based on the largest frequent patterns satisfying the support requirements. For instance, a common response to the information support strategy can be described by a slight decrease of the EDA, since values under 3 have a corresponding negative local angle. These results suggest that the alternative behaviors elicit different electrodermal responses. In particular, the first category, information support behaviour, had the desired effect of alleviating children’s stressful levels and potential negative affective state. The second category, tangible assistance, appears to be the strategy resulting in a more intense response (potentially revealing attention or stress) from the users. These results are in line with a previous study using the same scenario, in which we found that tangible assistance was the least preferred type of support by children [20]. The EDA levels increased possibly due to the tension reduction comments provided by the robot (e.g., “I always say, lucky in love, unlucky in chess”) that children found a bit too “embarrassing”.

The support behaviour categories, esteem and emotional support, were the ones who caused lower levels of EDA, suggesting that these behaviours are the ones with better efficacy on comforting the users (by putting them in a more relaxed stated). However, lower EDA levels are not always desirable: since high values are related with cognitive load, this can be a desired state in learning oriented scenarios.

C. Using EDA to Predict Children’s Affect

Similarly to the followed procedure for the analysis of responses to elicited supportive behaviour, the partitions framing the users’ affective state assume a tiny lag and a duration in agreement with the iCat predictions. 722 partitions were collected from the interaction logs of the 38 participants,

TABLE II: Questionnaire ratings for Engagement, Social Support, Help, Motivation, Co-Presence, Attentional Allocation, Perceived Message Understanding (PMU), Perceived Affective Understanding (PAU), Perceived Affective Interdependence (PAI) and Perceived Behavioural Interdependence (PBI).

	Engagement	Social Support	Help	Motivation	Co-Presence	Att. Allocation	PMU	PAU	PAI	PBI
First Interaction	4.42	3.75	4.21	4.81	4.69	4.65	4.50	4.27	4.27	4.15
Repeated Users	4.63	3.35	3.61	3.88	4.92	4.63	4.13	3.79	3.50	3.13
Easy Exercise	4.79	3.85	3.72	5.00	4.92	4.92	4.58	4.33	4.08	4.50
Difficult Exercise	4.50	3.83	4.67	4.79	4.43	4.29	4.21	4.64	4.57	4.71
All cases	4.56	3.65	4.03	4.54	4.75	4.62	4.34	4.20	4.05	3.99

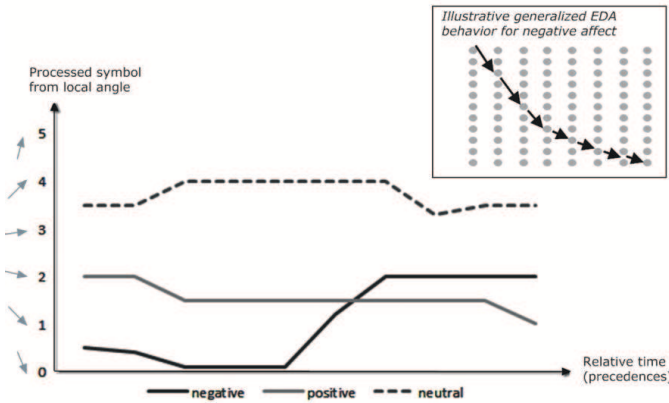


Fig. 6: Emerging sequence patterns for the different affective states of the children.

containing the probability of children’s negative, neutral and positive valence at different moments of the interaction obtained through the iCat’s affect recognition system. Again, coherent patterns emerged (from the application of sequential pattern miners over the pre-processed signal) across different affective states. Under an alphabet of 6-length, we were able to generalize the following sequences based on the largest frequent patterns for three major affective states: *i*) negative valence $\langle (1|0)(0)(0)(1|2|3)(1|2|3) \rangle$, *ii*) positive valence $\langle (2)(1|2)(1|2)(1|2)(1|2)(1) \rangle$, and *iii*) high score of neutral valence $\langle (3|4)(4)(4)(4)(3|4)(3|4) \rangle$ (see Fig. 6 for a graphical representation). We also considered the affective states where the positive valence score is repressed and where the negative valence score is repressed. As expected, the emerging patterns were, respectively, similar to the negative and positive affective states.

Interestingly, the states where the iCat predicts negative affective states, which commonly correspond to states where the user is neither moving nor conveying affective facial expressions, are precisely the states where EDA is higher. This seems to be correlated with states of high concentration, attention and reasoning. Additionally, negative states appear to be related with a local decrease of EDA baseline but a quick recovery of the level when compared against negative states.

We also carried out Pearson’s correlation tests between the affective states recorded in the interaction logs and the statistical and geometric metrics of the partitions for both

the unprocessed and the pre-processed versions of the signal. The following significant results were obtained with one-tailed probability. Negative valence is significantly correlated with quadratic error ($r=.09$, $p<.05$) in the unprocessed signal and with distortion in the unprocessed signal ($r=.10$, $p<.05$), although the correlation coefficients are not too strong. There was a significant relation between neutral valence and squared deviations ($r=-.07$, $p<.05$), product-moment ($r=.09$, $p<.05$) and kurtosis ($r=-.33$, $p<.001$) in the pre-processed signal. Neutral valence was also moderately correlated with kurtosis in the unprocessed signal ($r=.17$, $p<.001$). Finally, positive valence was significantly related to squared deviations ($r=.01$, $p<.05$) and quadratic error ($r=-.12$, $p<.05$) in the pre-processed signal, as well as to kurtosis in the unprocessed signal ($r=-.08$, $p<.05$).

Considering the results presented above, two main conclusions can be drawn:

- 1) The negative affective states are characterised by lower variations of the pre-processed signal (for being inversely correlated with dispersion features) but associated with increases in EDA, while the positive and negative affective states have more variations of the signal (being positively correlated with squared deviations and distortion, respectively), although these values were not captured by the pattern-view analysis.
- 2) The Pearson product-moment can be used to differentiate positive affective states from the remaining ones (negative and neutral), since this metric is negatively correlated with positive valence and positively correlated with the other two states. Similarly, low kurtosis (flat variations of the signal) can be used to exclude negative affective states.

VI. CONCLUSIONS AND IMPLICATIONS FOR FUTURE RESEARCH

In this paper, we presented one of the first HRI studies in which children’s electrodermal activity was measured in a real-world environment while interacting with a social robot. We started by describing the study protocol and how the collected data was processed and analysed. The results suggest that statistical measures applied to the signal collected during the whole interaction can be used to characterise children’s perception of the robot and the interaction. In particular, engagement, motivation and attention were found to be statis-

tically significant discriminated by electrodermal features. We also found that the different supportive behaviours employed by the robot (information support, tangible assistance, esteem and emotional support) can elicit different patterns in the EDA signal of the participants. Specific patterns also emerge when children are experiencing different affective states (negative, neutral or positive valence). Besides, significant correlations were found between statistical features of the signal and both the supportive behaviours and the valence of the children, suggesting that these metrics can help in the automatic prediction of children's motivational and affective states.

The results presented in this paper yield some interesting implications for future research in child-robot interaction. First, EDA data can be a valuable and reliable method for capturing children's interaction with social robots. As such, it can be used to complement and validate traditional survey methods such as questionnaires, interviews or video observation. The fact that the study was carried out in children's regular school environment may have contributed positively for this result since user studies in laboratory settings usually can cause higher levels of stress in the participants (which could undermine the EDA signals).

Another useful application of EDA data is on the automatic prediction of user's affective states. As reported in this paper, different patterns emerged when children were experiencing different affective states or were faced with different supportive behaviours displayed by the robot. Therefore, and considering that the sensors are becoming even more non-intrusive, we anticipate that EDA data will play a very important role in future affect recognition systems.

ACKNOWLEDGMENTS

The authors would like to thank to all the children from the elementary school in Parede, Portugal, who took part in the study. This research was supported by EU 7th FP under grant agreement no. 215554 and 317923, and by national funds through FCT-Fundação para a Ciência e a Tecnologia, under project PEst-OE/EEI/LA0021/2011, the PIDDAC Program funds and a PhD grant SFRH/BD/75924/2011.

REFERENCES

- [1] S. Ceci and M. Bruck, "Suggestibility of the child witness: A historical review and synthesis." *Psychological bulletin*, vol. 113, no. 3, p. 403, 1993.
- [2] M. H. Scullin and S. J. Ceci, "A suggestibility scale for children." *Personality and Individual Differences*, vol. 30, no. 5, pp. 843 – 856, 2001.
- [3] J. Read and S. MacFarlane, "Using the fun toolkit and other survey methods to gather opinions in child computer interaction," in *Proceedings of the 2006 conference on Interaction design and children*. ACM, 2006, pp. 81–88.
- [4] S. Jerritta, M. Murugappan, R. Nagarajan, and K. Wan, "Physiological signals based human emotion recognition: a review," in *Signal Processing and its Applications CSPA 2011 IEEE 7th International Colloquium on*. IEEE, 2011, pp. 410–415.
- [5] D. Fowles, "The measurement of electrodermal activity in children," *Developmental psychophysiology: theory, systems, and methods*, p. 286, 2008.
- [6] J. Wagner, J. Kim, and E. André, "From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification," in *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*. Ieee, 2005, pp. 940–943.

- [7] R. S. Aylett, G. Castellano, B. Raducanu, A. Paiva, and M. Hanheide, "Long-term socially perceptive and interactive robot companions: challenges and future perspectives," in *Proceedings of the 13th international conference on multimodal interfaces*, ser. ICMI '11. New York, NY, USA: ACM, 2011, pp. 323–326.
- [8] P. Rani, N. Sarkar, C. A. Smith, and L. D. Kirby, "Anxiety detecting robotic system - towards implicit human-robot collaboration," *Robotica*, vol. 22, no. 01, pp. 85–95, 2004.
- [9] D. Kulic and E. Croft, "Affective state estimation for human-robot interaction," *Robotics, IEEE Transactions on*, vol. 23, no. 5, pp. 991–1000, oct. 2007.
- [10] S. Y. Okita, V. Ng-Thow-Hing, and R. K. Sarvadevabhatla, "Multimodal approach to affective human-robot interaction design with children," *ACM Trans. Interact. Intell. Syst.*, vol. 1, no. 1, pp. 5:1–5:29, Oct. 2011.
- [11] K. Itoh, H. Miwa, Y. Nukariya, M. Zecca, H. Takanobu, S. Roccella, M. Carrozza, P. Dario, and A. Takanishi, "Development of a bioinstrumentation system in the interaction between a human and a robot," in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*. IEEE, 2006, pp. 2620–2625.
- [12] C. Liu, P. Rani, and N. Sarkar, "Affective state recognition and adaptation in human-robot interaction: A design approach," in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, oct. 2006, pp. 3099–3106.
- [13] E. Mower, M. P. Black, E. Flores, M. Williams, and S. Narayanan, "Rachel: Design of an emotionally targeted interactive agent for children with autism," in *Multimedia and Expo (ICME), 2011 IEEE International Conference on*, july 2011, pp. 1–6.
- [14] S. Reis and N. Correia, "An imaginary friend that connects with the user's emotions," in *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology*, ser. ACE '11. New York, NY, USA: ACM, 2011, pp. 1:1–1:8.
- [15] C.-C. Chang and C.-J. Lin, "Libsvm: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, May 2011.
- [16] G. Castellano, I. Leite, A. Pereira, C. Martinho, A. Paiva, and P. W. McOwan, "Inter-ACT: An affective and contextually rich multimodal video corpus for studying interaction with robots," in *Proceedings of the ACM International Conference on Multimedia*. ACM, 2010, pp. 1031–1034.
- [17] G. Castellano, I. Leite, A. Pereira, C. Martinho, A. Paiva, and P. McOwan, "Affect recognition for interactive companions: challenges and design in real world scenarios," *Journal on Multimodal User Interfaces*, vol. 3, pp. 89–98, 2010, 10.1007/s12193-009-0033-5.
- [18] C. Cutrona, J. Suhr, and R. MacFarlane, "Interpersonal transactions and the psychological sense of support," *Personal relationships and social support*, pp. 30–45, 1990.
- [19] I. Leite, A. Pereira, G. Castellano, S. Mascarenhas, C. Martinho, and A. Paiva, "Modelling empathy in social robotic companions," in *Proceedings of the 19th international conference on Advances in User Modeling*, ser. UMAP'11. Berlin, Heidelberg: Springer-Verlag, 2012, pp. 135–147.
- [20] I. Leite, G. Castellano, A. Pereira, C. Martinho, and A. Paiva, "Long-term interactions with empathic robots: Evaluating perceived support in children," in *Proceedings of the International Conference on Social Robotics*, ser. to appear. Springer, 2012.
- [21] C. Sidner, C. Kidd, C. Lee, and N. Lesh, "Where to look: a study of human-robot engagement," in *Proceedings of the 9th international conference on Intelligent user interfaces*. ACM, 2004, pp. 78–84.
- [22] A. Gordon, "Assessing social support in children: Development and initial validation of the social support questionnaire for children," Ph.D. dissertation, Faculty of Louisiana State University and Agriculture and Mechanical College, University of Houston, 2011.
- [23] I. Leite, S. Mascarenhas, A. Pereira, C. Martinho, R. Prada, and A. Paiva, "“why can't we be friends?” an empathic game companion for long-term interaction," in *Proceedings of the 10th international conference on Intelligent virtual agents*, ser. IVA'10. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 315–321.
- [24] F. Biocca, "The cyborg's dilemma: Embodiment in virtual environments," in *Cognitive Technology, 1997. Humanizing the Information Age. Proceedings., Second International Conference on*. IEEE, 1997, pp. 12–26.
- [25] C. Antunes and A. L. Oliveira, "Mining patterns using relaxations of user defined constraints," in *Proc. of the Workshop on Knowledge Discovery in Inductive Databases*, 2004.