# Physiological Detection of Affective States in Children with Autism Spectrum Disorder

Sarah Sarabadani, Larissa C. Schudlo, Ali Akbar Samadani, *Member, IEEE*, and Azadeh Kushski, *Member, IEEE*

**Abstract**—Autism spectrum disorder (ASD) is associated with emotion processing difficulties, including limitations in understanding the emotional states of others and processing one's own internal experiences. The nature of these difficulties remains largely unknown. This is due, in part, to challenges in acquiring reliable self-reports of emotional experiences from this population. Automatically characterizing emotional states with the use of physiological signals is a potential means of overcoming this problem, as physiological signals can provide an objective and nonverbal method for assessing affective states. However, this approach has not been well considered with ASD to date. To this end, we investigated detection of autonomic responses to positive and negative stimuli in children with ASD using four physiological measurements. Electrocardiograms, respiration, skin conductance and temperature were measured while 15 children with ASD viewed standard images known to evoke varying levels of valence (positive and negative) and arousal (low and high intensity). Using an ensemble of classifiers, affective states induced by stimuli of positive and negative valence or high and low arousal was differentiated at average accuracies approaching or exceeding 80 percent. These results suggest the feasibility of discerning affective states in individuals with ASD objectively using physiological signals.

**Index Terms**—Emotion detection, physiological signals, autism spectrum disorder, classification

## 1 Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by social communication difficulties, and repetitive and restricted behaviors and interests [1]. Social impairments in this population are often linked to emotional processing deficits [2]. An extensive body of literature on this topic suggests that individuals with ASD often experience difficulty ascertaining their own mental and emotional states as well as those of others [3], [4], [5], [6]. It has been reported that half the population with ASD is affected by alexithymia (difficulties in discerning and describing one's internal body state) [7], whereas the occurrence in the typically-developed population is only one in ten [8], [9]. Lambie and Marcel characterized the emotional experience in a two-level model [10]. A first order experience is one that induces neurophysiological arousal consequent of an emotional state. The second order experience is the self-awareness of this emotion (introception).

Despite reduced self-reported scores of emotional awareness, individuals with ASD do not necssarily exhibit a diminished response in brain regions underpinning the first order emotional experience (*i.e.*, the amygdala-orbitofrontal system) [6]. This disconnect between perceived emotions and measured functional response suggests that alexithymia symptoms in ASD do not necessarily stem from the brain's inability to respond to emotions. On the other hand, individuals with ASD tend to exhibit impairments in second-order emotional experiences (*i.e.*, awareness of bodily states) [9], [11], [12], [13]. A negative correlation between activity levels in the insula, a key region implicated in interoception, and the degree of alexithymia in ASD has been observed [6]. The apparent disassociation between emotional arousal and conscious awareness of this response may be an underlying factor of alexithymia in ASD. Challenges in emotional processing have also been linked to increased severity of psychiatric disorders such as depression [2].

A large body of literature has considered the interplay between one's emotional state and resulting autonomic activity in the typically developing/ed population (Table 1). Changes in one's affective state can elicit a notable response in their autonomic nervous system (ANS) [14]. This response, in turn, can be detected and quantified with physiological measurements such as heart rate and respiration [15], [16]. For example, anger is often accompanied by increases in heart rate and blood pressure [17]. To explore emotion-specific responses of the ANS, Ekman et al. considered cardiac activity, skin temperature and resistance and

- *S. Sarabadani, L. C. Schudlo, A. A. Samadani and A. Kushki are with the Holland Bloorview Research Insititue, Holland Bloorview Kids Rehabiliation Hospital, 150 Kilgour Road, Toronto, ON M4G 1R8, Canada. E-mail: sarah.sarabadani@mail.utoronto.ca, lschudlo@ryerson.ca, a.a.samadani@gmail.com, akushki@hollandbloorview.ca.*
- *S. Sarabadani, A. A. Samadani and A. Kushki are with the Institute of Biomaterials & Biomedical Engineering, University of Toronto, Toronto, ON M5S 3G9, Canada. E-mail: sarah.sarabadani@mail.utoronto.ca, a.a.samadani@gmail.com, akushki@hollandbloorview.ca.*
- *L. C. Schudlo is with the Department of Electrical and Computer Engineering, Ryerson University and a member of the Institute of Biomedical Engineering, Science and Technology (iBEST) at Ryerson University and St. Michael's Hospital, 350 Victoria Street, Toronto, Ontario M5B 2K3, Canada. E-mail: lschudlo@ryerson.ca.*

*Manuscript received 1 June 2017; revised 5 Mar. 2018; accepted 18 Mar. 2018. Date of publication 27 Mar. 2018; date of current version 25 Nov. 2020. (Corresponding author: Larissa C. Schudlo).*
*Recommended for acceptance by W. Zheng.*
*Digital Object Identifier no. 10.1109/TAFFC.2018.2820049*

TABLE 1
Previous Studies on Examining Emotion Recognition in Typically Developing or Developed Individuals via Autonomic Response

| Reference | Emotions | # of par | Elicitation method | Signals | Classifier | Accuracy (%) |
|---|---|---|---|---|---|---|
| [19] | Sad, anger, stress, surprise | 125 | Multimodal | ECG, Temperature, SC | SVM | 78.4 (3 emotions)<br>61.8 (4 emotions) |
| [20] | Joy, Anger, Sad, Pleasure | 3 | Music | EMG, ECG, SC, Respiration | LDA | 95 (personal)<br>70 (group) |
| [21] | Valance, Arousal | 36 | Robot actions | SC ECG | HMM | 83 Arousal ,80 Valance (personal ) 66 Arousal , 66 Valance (group) |
| [22] | Neutral, Anger, Hate, Grief, Platonic love Romantic love, Joy, Reverence | 1 | Personalized imagery | EMG, BVP, SC, Respiration | Hybrid LDA | 81(personal) |
| [23] | Happiness, Disgust, Fear | 9 | IAPS pictures | EMG, ECG, SC, Respiration | KNN<br>Random Forest | 62.70 (group)<br>62.41(group) |
| [24] | Valance, Arousal | | IAPS pictures | EMG, ECG, SC, Temperature, BVP , Respiration | Neural Networks | Valance 89.7 (personal) Arousal 63.76 (personal) |
| [25] | Sad, Anger, Surprise, Fear, Frustration, Amusement | 14 | Movies | SC, ECG | KNN, DFA Marquardt Back Propagation | 71(personal):KNN<br>74(personal):DFA<br>83(personal):MBP |
| [26] | Joy, Anger, Sad, Pleasure | 1 | Music | ECG, EMG, SC, Respiration | SVM | 76 (Fission ,personal) and 62 (Fusion, personal) |
| [27] | Joy, Anger, Sad, Pleasure | 1 | Music | EMG | Neural Network | 82.29 (personal) |
| [28] | Joy, Anger, Sad, Pleasure | 1 | Music | ECG, EMG, SC, Respiration | LDA | 83.4 (personal) |
| [29] | Active and passive arousal/ Valence | 44 | IAPS and video game | ECG | LDA | 76.19 (personal) |
| [30] | Engagement, Confusion, Boredom, Hopefulness | 1 | Studying | SC, BVP, EEG | KNN, SVM | Best result: 86.3 |
| [31] | Boredom, Engagement, Anxiety | 20 | Tetris game | SC, BVP, Respiration, Temperature, EEG | LDA, QDA, SVM | 63 (personal) |
| [32] | Stress, Relaxed | 32 | Paced Stroop test | SC, BVP, Temperature, Pupil diameter | SVM | Best result: 90.10 |
| [33] | Arousal, Valence, Liking, Dominance, Familiarity | 32 | Music | SC, EMG, EOG, BVP, EEG, Temperature, Respiration | Fisher's Linear Discriminant, Naïve Bayes | Best result: 65.1 |
| [34] | Frustration | 36 | Videos Computer game | SC, BVP | Hidden Markov Model | 67.4 (personal) |

*BVP = blood volume pulse, ECG = electrocardiogram, EEG = electroencephalography, EMG = electromyography, SC = skin conductance, SVM = support vector machine.*

muscle tension during 6 emotion states: surprise, disgust, sadness, anger, fear and happiness [18]. Physiological differences were detected between positive and negative emotional states, and even some of the negative states were discernable. Similarly, Picard et al. differentiated eight discrete emotions (neutral, anger, hate, grief, platonic love romantic love, joy, and reverence) using measures of muscle tension, heart activity, skin conductance, and respiration [22]. Using the same four physiological measures, Kim and André found distinguishable patterns between states of positive/high arousal, negative/high arousal, positive/low arousal, and negative/low arousal [20]. Kim K. H. et al. has also investigated autonomic patterns in heart rate, skin conductance, and temperature associated with states of sadness, anger, stress, and surprise [35]. These major findings are only a fraction of the investigations that confirm the

existence of identifiable physiological patterns that distinguish affective states.

Although emotional induction and autonomic responses have been well considered in the typically developed population, this area has not been well explored in ASD. Indeed, changes in emotional states can elicit physiological arousal in individuals with ASD [11]. However, patterns may be atypical [22]. Groden et al. [23] and Ben Shalom et al. [11] investigated physiological reactivity in response to emotional stimuli in ASD and were able to detect notable differences among various arousal conditions. In particular, Groden et al. considered heart rate variations in response to four different stress-inducing conditions. Atypically high basal heartrates were observed, as well as high inter-subject variability in response to the stress conditions. Using skin conductance, Ben Shalom et al. investigated the response to pleasant, unpleasant, and neutral stimuli in children with ASD. Physiological responses measured from children with ASD were not significantly different from typically developing children, despite differences self-reports ratings. Facial expression, characterized by facial electromyography (EMG), in response to viewing dynamic audio-visual displays of happiness, anger, and fear was characterized in individuals with ASD and typically developing individuals by Rozga et al. [24]. Typically developing individuals demonstrated different responses to positive and negative stimuli, whereas individuals with ASD showed no significant differences across the emotions. Yet, behavioral performance in identifying emotions (accuracy and speed) did not vary across the two groups. These results support the potential of using physiological measures as more sensitive means of identifying differences in emotion-selective responses across ASD and typically developing individuals than behavioral performance. These works considered across-group statistical analyses, in which specific variables were compared for mean differences across conditions or populations.

A single-trial analysis may reveal more information and better elucidate the mechanisms of ASD than summarizing data into a mean value for a group-level analysis of this highly heterogeneous condition. A single-trial analysis may be of particular value when considering emotional or affective responses, which are highly individualized [25] and variable within a subject [26]. A single-trial, rather than group-averaged approach, may permit the discovery of individualized, subject-specific patterns not necessarily identifiable by traditional statistics, which often rely on assumptions of normality and linearity. To our knowledge, only two papers have performed automatic classification of emotional states on a single-trial basis in ASD. Liu et al. classified affective states of anxiety, engagement, and liking at average accuracies of 79.5 percent, 84.3 percent, and 85.0 percent, respectively, using a number of measures (ECG, electro dermal activity (EDA), electromyogram (EMG), blood volume pulse (BVP), temperature, bio impedance, and heart sound) in children with ASD [27]. Supervised classification was performed using a support vector machine trained and tested on data samples labeled for affective state by a behavioral therapist. Similarly, Kushki et al. classified anxiety-related arousal at an average accuracy of 95 percent using cardiac activity (R-R interval duration) in an unsupervised learning approach based on a modified-Kalman filter [22]. Emotional stimuli can be characterized by dimensions of valence (positive *vs* negative) and arousal (intense *vs* neutral). While majority of investigations have considered arousal or discrete emotional states in the ASD population, no previous work has considered both arousal and valence concurrently. To this end, we sought to automatically characterize affective states along axes of both valence and arousal using physiological measures in children with ASD. Specifically, we sought to evaluate the classification accuracies achievable in automatically differentiating affective states of: i) positive *vs* negative valence during high arousal and ii) positive *vs* negative valence during low arousal using physiological measurements of ECG, EDA, respiration and temperature. A reliable, physiologically-based method of detecting affective states in real-time could provide a language-free, non-invasive and economically-feasible means of recognizing body states for individuals with ASD. This can be used in therapeutic interventions or affective-sensitive assistive devices that aid individuals with ASD in recognizing and/or communicating their internal states [25], [27]. Our primary focus was to evaluate the feasibility of developing such a technology, rather than the studying the biological underpinnings of emotional responses in ASD. This study represents a first step toward developing a reliable measure of affective states in ASD that could be used in real-time to automatically identify and address changes in one's internal state in everyday situations.

## 2 Methods

### 2.1 Participants

Fifteen children (3 female) with ASD between 12 and 18 years were recruited from Holland Bloorview Kids Rehabilitation Hospital to participate in this study. Participants had a clinical diagnosis of ASD using DSM-IV criteria verified by the Autism Diagnostic Observation Schedule [39] and the Autism Diagnostic Interview - revised (ADI-R). To assess intelligence, ASD symptomatology, and related comorbidities, participants completed the Weschler Abbreviated Scale of Intelligence, the Social Communication Questionnaire (SCQ), and the Child Behaviour Checklist (CBCL), respectively. These were administered by research-reliable experts at the Province of Ontario Neurodevelopmental Disorders (POND) network. All participants had a full-scale IQ score above 70, and no sensory impairments such as deafness or blindness.

Ethics approval for this study was obtained from the Bloorview Research Institute and the University of Toronto research ethics boards. Written consent was obtained from all participants in the study (*i.e.*, both the child and parent).

For children who did not have the capacity to consent, their parents consented on their behalf following their assent.

### 2.2 Instrumentation

Four physiological signals, namely electrocardiogram (ECG), skin conductance (SC), respiration, and skin temperature, were recorded using the Procomp Infiniti system (Thought Technology Ltd). Cardiac activity was measured using three leads secured to the participant's torso. SC was measured using a pair of dry Ag-AgCl electrodes secured to the palmar surface of the proximal phalanges of the second and third digits of the participant's non-dominant hand. Respiration was measured using a piezoelectric belt secured

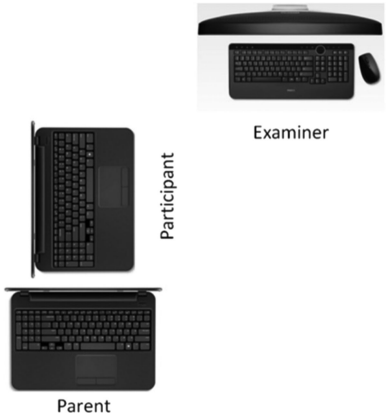


Fig. 1. Experimental setup. Participants viewed images and completed SAM on one laptop. Parents faced their child and completed the SAM on another laptop. The experimenter was seated at a computer behind the participant.

around the abdomen. Lastly, skin temperature was recorded using a thermistor placed on palmar surface of the distal phalanx of the fourth digit of the hand. ECG data were recorded at frequency of 2048Hz. All other measures were recorded at frequency of 256Hz.

The layout of the experiment room is shown in Fig. 1. The participant was positioned in front of one laptop, on which they viewed the emotional stimuli and rated their affective state. The child's parent was seated in front of a second laptop such that they could see their child and rate their child's emotional reactions, but not the emotional stimuli the child observed. The experimenter was seated at the desktop behind the participant.

### 2.3 Stimuli

The International Affective Picture System (IAPS) is the gold standard method of eliciting various levels of arousal and valence [23], [24], [40], [41], [30]. This database consists of 956 images with a variety of culturally non-specific themes. Similar to the IAPS, the Geneva Affective Picture Database (GAPED) is a newer image database that contains 730 images. Images from both the IAPS and GAPED were used in this study.

For every picture, both databases report ratings on three scales: arousal, valance, and dominance. In the IAPS, each picture has a rating between 1 and 9 (1: lowest arousal/pleasure/dominance, 9 the highest arousal/pleasure/dominance). Each picture in the GAPED is rated between 0 and 100 (where 100 is the highest arousal, most positive) [42]. Images in IAPS were evaluated using the self-assessment manikin (SAM). This same scale was used by participants in this study.

IAPS and/or SAM has been employed in several studies involving children with ASD (*e.g.*, [11], [6], [43]). GAPED has been employed in studies considering typically developing individuals [44], but has not been considered with the ASD population.

Images depicting four combinations of arousal and valance were considered in this work: high arousal/positive valence, low arousal/positive valence, high arousal/negative valence, and low arousal/negative valence. Images above a mid-scale rating (*i.e.*, 4.5 for IAPS, and 50 for GAPED) were considered high in arousal or positive in valence dimensions. Images below a mid-scale rating were considered low in the given dimension.

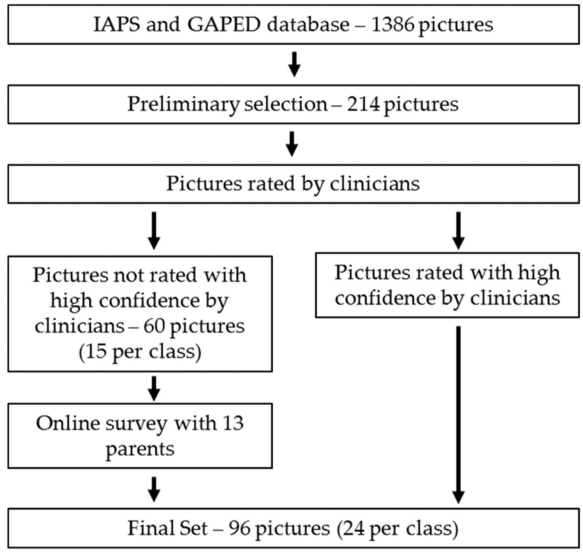


Fig. 2. Procedure for picture selection.

The image selection procedure is shown in Fig. 2. A total number of 214 pictures were selected from the databases. Pictures of faces, erotic photos, and those illustrating savagery and mutilation themes were deemed unsuitable for the age range of the participants and were excluded. The remaining images were then evaluated by clinicians. After the clinician refinement process, 60 of the images (15 per theme) were not rated with high confidence. To refine this remaining set of images, 13 parents evaluated and commented on the 60 pictures via an online survey. They indicated whether each image elicited positive or negative emotions, and if the response was weak or strong. The final set of 96 images (24 per type) consisted of images selected in the preliminary screening combine with parents' choices. The average of the actual ratings of the final set of images is shown in Table 2.

### 2.4 Experimental Protocol

At the start of the session, the experimental paradigm was explained to both the participant and their parent. To ensure an understanding of the SAM, participants were asked to rate two practice images (an snake and ice cream) using the SAM and to verbalize their feeling about the images, to ensure they aligned with their ratings. Participants were instructed to remain as still as possible while viewing images to mitigate any movement-induced artifacts. Sensors were affixed to the child and they completed a practice trial to ensure their understanding of the study. Participants were given breaks throughout the session as needed.

TABLE 2
Average (Mean ± Standard Deviation) of the Actual Ratings of Final Pictures Selected

| | Valence | Arousal |
|---|---|---|
| HP | 6.91±0.44 | 5.38±0.98 |
| HN | 3.67±0.53 | 5.98±0.67 |
| LP | 8.1±0.45 | 2.16±1.38 |
| LN | 3.41±1.2 | 5.11±0.98 |

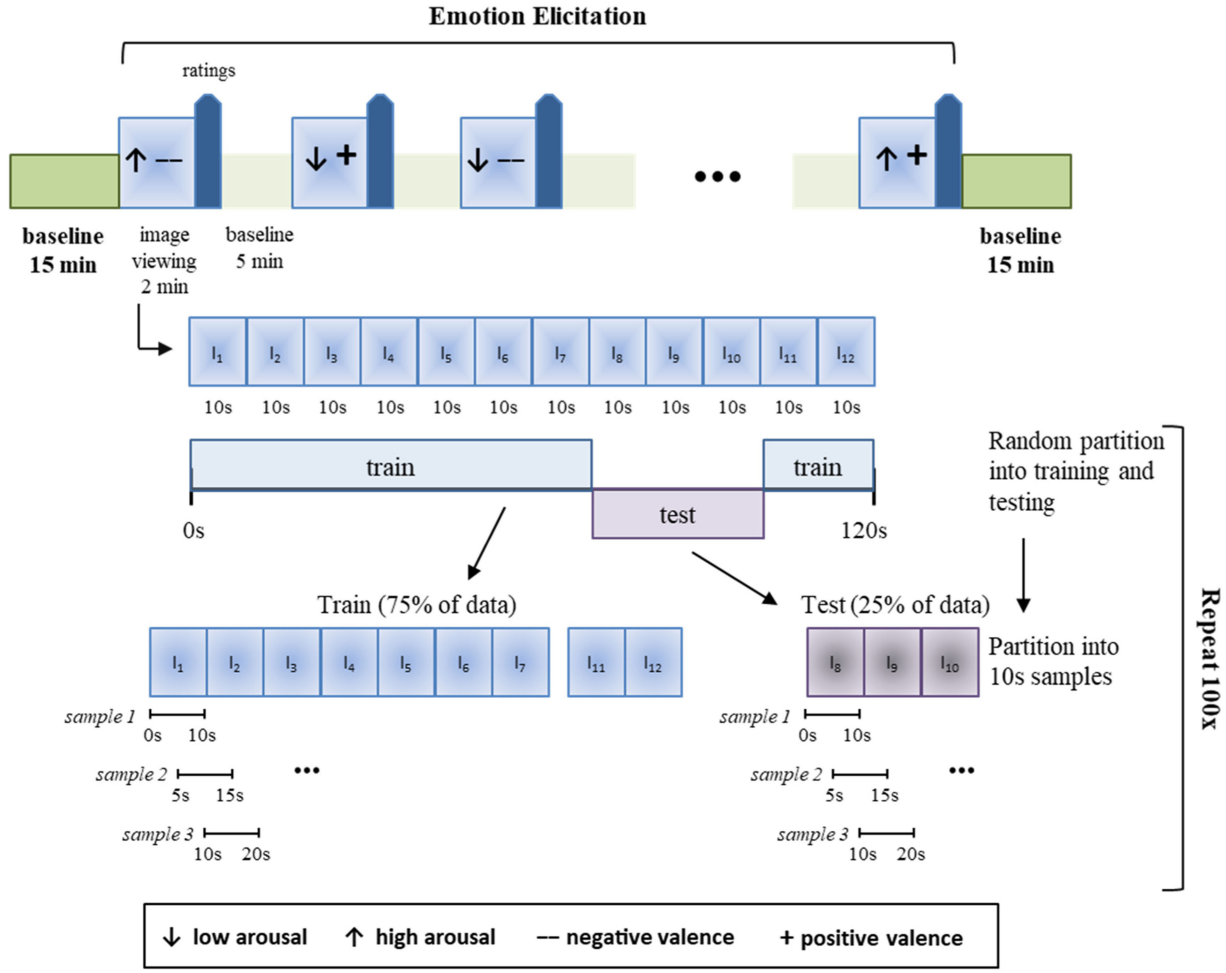


Fig. 3. Timing paradigm of experimental protocol.

Fig. 3 shows the timing paradigm of each experimental session. The session began and concluded with a 15-minute baseline interval, during which participants viewed an animated movie. This time allowed participants to acclimate to the experimental environment. The initial baseline period was followed by 2 minute blocks of stimulus presentation, interspersed with 5-minute baseline periods. During the blocks of stimulus presentation, participants viewed 12 images for 10s each. Each block was intended to evoke one of the four affective states (*i.e.*, HP, HN, LP, or LN). Two blocks of each state were collected, yielding 4 minutes of physiological data per affective state. The order of stimulus block types, the images used in each block and the order of images within each block was random for each participant. Following each block, both the child and parent completed the SAM. Parents were blind to the images their child viewed and were instructed to evaluate the emotional response they observed from their child.

## 3 Analysis

### 3.1 Preprocessing

The raw data were visually inspected to identify any significant movement-induced artifacts. Majority of the data were free of artifacts, with no more than approximately 10 percent of the data of a given physiological signal from a given participant being contaminated with artifacts. Thus, all data were considered for analysis.

*ECG:* The data were bandpass filtered using an 5th order type II Chebyshev filter with low and high cutoff frequencies of 0.5 Hz and 25 Hz, respectively, to attenuate noise stemming from powerline noise, baseline drift, and physical movement. R-peaks in the ECG were determined using the algorithm presented in [45] via the BioSig Matlab toolbox [46]. Visual inspection of randomly selected data confirmed the accuracy of the algorithm. Instantaneous heart rate was determined by finding the inverse of the R-R interval durations. A ninth order median filter (*i.e.*, considered 9 points at a time) was then applied to the detected R-peaks to remove any outliers.

*Respiration:* Data were filtered using a 5th order low-pass type II Chebyshev filter with a cut off frequency of 0.5 Hz. This frequencywere selected given that the shortest breathing interval in adults, on average, is 3 seconds [47]. Peaks in the respiration signal were identified using the criterion that peaks would fall at least 1s apart to account for increased respiration due to arousal. Intervals less than 3 seconds apart were confirmed with visual inspection.

*Skin conductance:* To minimize baseline drifts consequent of thermoregulation and perspiration-induced changes in sensor adhesion, measurements were first detrended over the entire session. The signal was then low-pass filtered using a $4^{th}$ order Butterworth filter with a 1Hz cut-off frequency.

*Temperature:* The signal was first detrended and then filtered using a $4^{th}$ order low-pass Butterworth filter with cut off frequency 1 Hz.

TABLE 3
List of Features

| Signal | No. | Feature |
|---|---|---|
| ECG | 1 | Mean RR interval |
| | 2 | Minimum RR interval |
| | 3 | Maximum RR interval |
| | 4 | Standard deviation of RR intervals |
| | 5 | Mean heart rate |
| | 6 | Minimum heart rate |
| | 7 | Maximum heart rate |
| | 8 | Standard deviation of heart rates |
| | 9 | Median of top quartile of heart rates |
| | 10 | Median of bottom quartile of heart rates |
| | 11 | Slope of the instantaneous heart rate |
| SC | 12 | Mean SC |
| | 13 | Slope of SC signal |
| Temperature | 14 | Mean temperature |
| | 15 | Standard deviation of temperatures |
| | 16 | Minimum temperature |
| | 17 | Maximum temperature |
| | 18 | Slope of temperature signal |
| Respiration | 19 | Mean respiration interval |
| | 20 | Minimum respiration interval |
| | 21 | Maximum respiration interval |
| | 22 | Minimum respiration rate |
| | 23 | Maximum respiration rate |

*SC: Skin Conductance, ECG: Electrocardiogram*

Prior to feature extraction, the average of the last two minutes of the previous baseline was subtracted from each task block for each of the four signals to mitigate any carry-over effects from the previous task.

### 3.2 Feature Extraction

Analyses were performed offline using cross-validation. The data from each block were randomly partitioned into training and testing segments (Fig. 3). Seventy-five percent of the data were used from training, and the remaining 25 percent were used for testing. The training and testing segments were then parsed into 10-second sub-windows with 5 seconds of overlap between adjacent sub-windows. Each of these sub windows was considered a sample and features were extracted from each sample. Therefore, a sample consisted of a 10s interval during which the participant viewed a single image, or a 10s interval during which the participants viewed 2 images (each for 5s) of the same type of stimuli. Training data were used only for feature selection and classifier training, while testing data were only used for classifier testing. No samples with a window overlapping both the training and testing segments were included in a given run of classification.

Table 3 summarizes features extracted from each sensor. Although frequency-based features are commonly extracted from ECG data, the short duration of each sub window considered for feature extraction rendered features of this type impractical. Thus, only time-domain ECG features were considered in this work.

### 3.3 Feature Selection

Feature selection was performed using a forward floating sequential search [48] with the Fisher Criterion as the selection criterion. A 3-dimensional feature set was selected from the candidate feature pool and used for classifier training.

TABLE 4
Demographic Information

| Measure | Value |
|---|---|
| Age (years) | 14.9±1.8 |
| Sex (Male:Female) | 12:3 |
| SCQ Score | 19.9±5.8 |
| Full-Scale IQ | 99.6±19.4 |
| Medication (Yes:No) | 7:8 |
| CBCL (Internalizing Problems) | 61.4±7.2 |

### 3.4 Classification

Two classification problems were addressed: 1) high arousal/negative valence vs high arousal/positive valence, and 2) low arousal/negative valence vs low arousal/positive valence. Given the challenges with self-evaluation of emotions in ASD, the actual image labels from the standardized databased from which they came were used to label the the data for classification. The valence and arousal ratings of the stimuli from standardized image databases have been validated in neurotypical populations, and the standardized images therefore should effectively elicit the intended affective reactions. The image ratings have not been validated with a large number of individuals with ASD.

Five different classifiers were considered: K-nearest neighbours (KNN) with $\mathrm{k} = 3$, linear discriminant analysis (LDA), and a support vector machines using linear, polynomial and RBF kernels. Both linear and nonlinear classi-fier varieties have been used in previous studies considering emotion classification. In addition to considering classifiers individually, classifier outputs were combined via majority vote (MV) to potentially enhance accuracies.

### 3.5 Performance Evaluation

The data were randomly partitioned into training and testing sets 100 times to evaluate classifier accuracy:

$$Accury(\%) = 100 * (N_{correct}/N_{test}). \quad (1)$$

where $N_{correct}$ denotes the number of correctly classified test points and $N_{test}$ denotes the total number of testing samples. The mean accuracy across all runs of training/testing partitioning was determined to yield an average classification accuracy for each participant. To evaluate if classification accuracies were greater than random chance (*i.e.*, 50 percent for a binary problem), classification using the majority vote classification scheme was repeated using randomized data labels.

## 4 Results

### 4.1 Participant Demographics

Data from one participant were excluded prior to analysis due to technical issues with the instrumentation. The demographic information for the remaining participants is shown in Table 4.

### 4.2 Classification Results

Average individual- and across-participant classification accuracies obtained with each classifier in classifying high

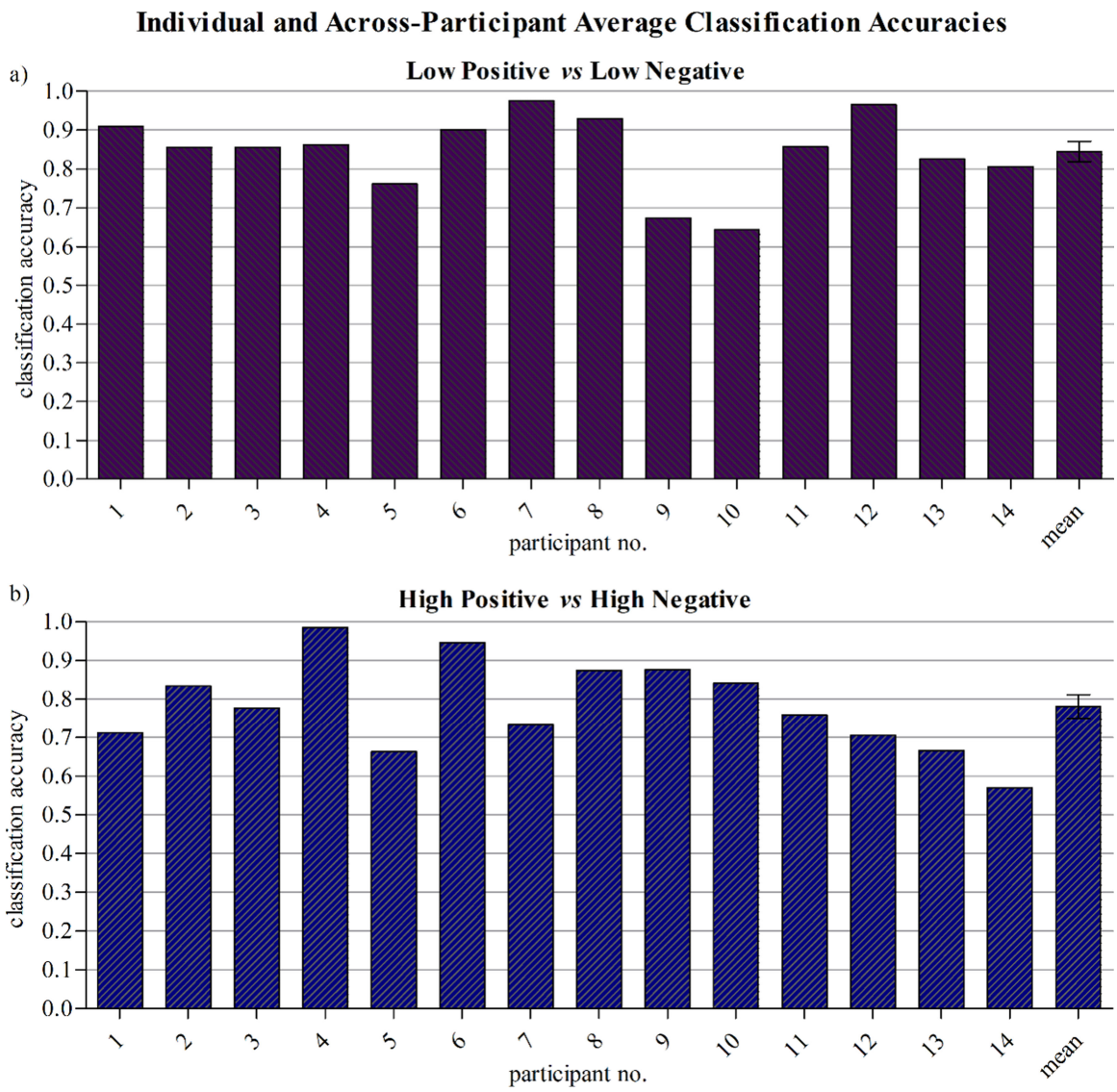


Fig. 4. Average individual and across-participant classification accuracies for classifying a) low positive vs low negative images and b) high positive vs high negative images using majority vote of 5 classifiers. Error bars represent standard error the mean.

arousal/positive valence *vs* high arousal/negative valence, and low arousal/positive valence *vs* low arousal/negative valence are shown in Fig. 4.

The highest across-participant average accuracies were achieved using a majority vote of all 5 classifiers, with accuracies of $84.5 \pm 9.8$ percent and $78.1 \pm 11.7$ percent obtained in differentiating low/positive *vs*. low/negative and high/positive *vs*. high negative, respectively. The average accuracy in differentiating low/positive *vs*. low/negative using a majority vote of the 5 classifiers is not greater than that obtained using a single, linear discriminant classifier ($83.5 \pm 11.8$ percent).

For a binary problem considering 92 samples, an accuracy equal or greater than 58.70 percent is considered statistically greater than chance for a binomial distribution (for $\alpha = 0.05$) [49]. Given this consideration, all individual participant classification accuracies obtained using the majority vote of the 5 classifiers were greater than chance levels. Table 5 shows the classification results obtained using the majority vote classifier with randomized labels. The data samples were labelled randomly, rather than using the correct affective states, and subject to classification with these labels (as in Sections 3.3-3.5). Classification accuracies using correctly-labelled participant data and randomized

TABLE 5
Classification Results Using Randomized Labels

| Participant | LP vs. LN | HP vs. HN |
|---|---|---|
| 1 | 0.490 | 0.497 |
| 2 | 0.496 | 0.507 |
| 3 | 0.496 | 0.483 |
| 4 | 0.505 | 0.504 |
| 5 | 0.498 | 0.491 |
| 6 | 0.501 | 0.504 |
| 7 | 0.497 | 0.484 |
| 8 | 0.515 | 0.504 |
| 9 | 0.497 | 0.488 |
| 10 | 0.495 | 0.498 |
| 12 | 0.507 | 0.513 |
| 13 | 0.498 | 0.510 |
| 14 | 0.509 | 0.510 |
| **Mean** | $\mathbf{50.0 \pm 0.7}$ | $\mathbf{49.9 \pm 1.0}$ |

Table 6
Confusion Matrix of Ensemble of Methods
(Across All Participants and Iterations of Cross-Validation)

| | | Predicted labels | |
|---|---|---|---|
| | **N = 28000** | **Low/Negative** | **Low/Positive** |
| **Actual labels** | **Low/Negative** | 11726 | 2077 |
| | **Low/Positive** | 2274 | 11923 |

Table 7
Confusion Matrix of Ensemble of Methods
(Across All Participants and Iterations of Cross-Validation)

| | | Predicted labels | |
|---|---|---|---|
| | **N = 28000** | **Low/Negative** | **Low/Positive** |
| **Actual labels** | **High/Negative** | 10852 | 2982 |
| | **High/Positive** | 3148 | 11018 |

labels would be similar if accuracies achieved using correct labels were due to random chance or erroneous processing. As expected, accuracies achieved with randomized labels are at levels of random chance ($\sim$50 percent), affirming the validity of the binomial test and our actual classification accuracies.

Tables 6 and 7 show the confusion matrices for classification results obtained using the majority vote classification scheme. Across all participants, misclassification rates were comparable across the two classes for each classification problem. However, differentiation of positive and negative valence during high arousal yielded greater misclassifications than low arousal.

### 4.3 SAM Results

Fig. 5 shows the relation between classification accuracies and accuracy of the participant ratings of the images considered in the given classification problem (*i.e.*, the agreement between participants' SAM ratings and actual ratings of the images).

Fig. 6 shows the accuracy of participant and parent ratings in comparison to the actual image ratings, and the agreement between participant and parent ratings for all images. No participants rated the images such that their ratings fully agreed with actual ratings. Five participants (1, 2, 8, 12, and 13) only selected one type of arousal for all images (*i.e.*, either all high or all low). Consequently, classification with these labels would be impossible. Similarly, three participants (1, 9 and 12) categorized all images with the same valence type.

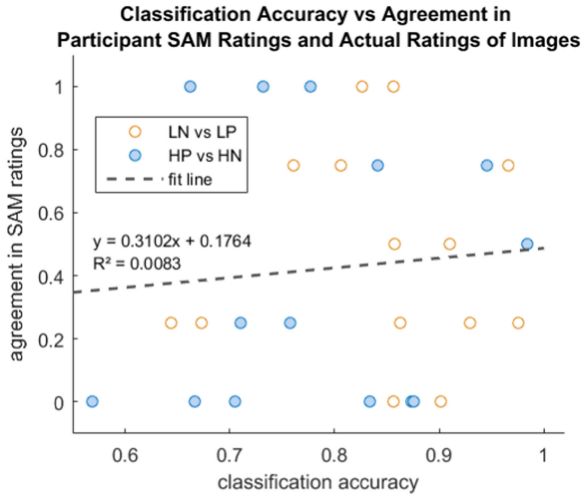


Fig. 5. Agreement between actual ratings and a) parent ratings and b) participant ratings of images using SAM.

### 4.4 Feature Selection

The five most frequently selected features using the majority voting classification scheme are shown in Table 8 for each classification problem.

Figs. 7 and 8 are heat maps of selected features for each participant. The colour intensities denote the number of times each feature was selected across all runs of classification.

## 5 Discussion

### 5.1 Classification

Results of this study indicate that affective states elicited by viewing positively and negatively valanced images can be

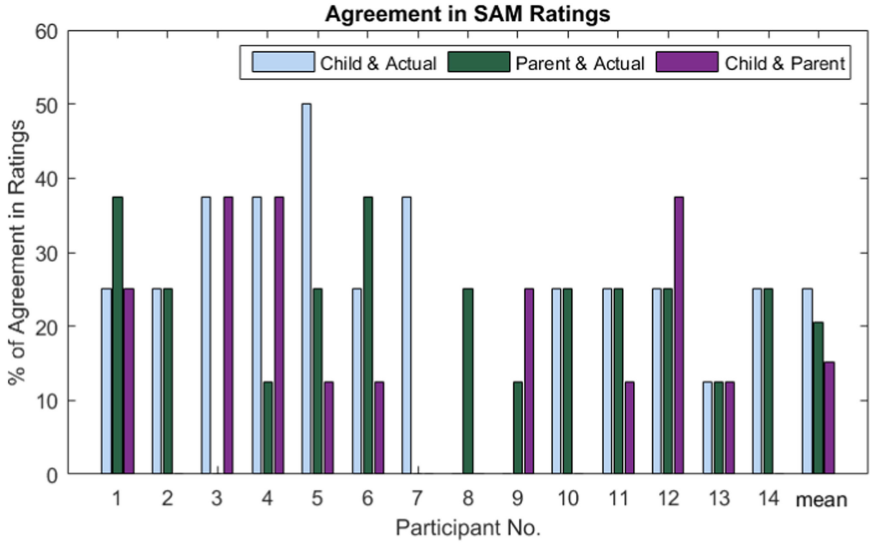


Fig. 6. Child's rating on high and low arousal images and positively and negatively valenced images. The horizontal black bar depicts the expected selection of each arousal type (*i.e.*, 4 images per type).

TABLE 8
5 Most Frequently Selected Features

| Feature Ranking | HP vs. HN | LP vs. LN |
|---|---|---|
| 1 | Mean SC | Mean SC |
| 2 | Minimum temperature | Minimum temperature |
| 3 | Maximum temperature | Maximum temperature |
| 4 | Maximum RR interval | Standard deviation of temperature |
| 5 | Maximum heart rate | Slope of temperature |

classified automatically using physiological measurements at average accuracies near 80 percent under both high and low arousal states in children with ASD.

Differentiation of valences was more accurate during a state of high arousal. The ratings of the stimuli suggest that this discrepancy may be attributed to the potency of the selected stimuli within and across the classification problems. The average valence levels were substantially different for positive and negative states within each classification problem, and were fairly consistent across the two problems (see Table 2). However, the arousal ratings were rated closer to neutral for three of the four sets of images. This was due to the limitations in images that could be used (due to ethical and age considerations, see Section 2.3 *Stimuli*), and may have contributed, in part, to differences in accuracies obtained for the two binary classification problems considered in this work.

Nonetheless, average accuracies of 84.5 percent and 78.1 percent were achieved in differentiating low/positive *vs.* low/negative and high/positive *vs.* high negative respectively, using an ensemble of classifiers. Although our results suggest similar average accuracies can be achieved using a single linear discriminant classifier in differentiating lo/positive *vs* low/negative states, not all participants exceeded levels of random chance with this classifier by the binomial test (P9 was lower than chance levels). Perhaps other ensemble learning techniques, such as bagging or boosting, would further enhance results and could be considered in future work. The confusion matrices in Tables 6 and 7 indicate that, on average, the classifier did not favour one class over the other. We were able to discriminate each pair of affective states equally with high accuracy for both classification problems considered.

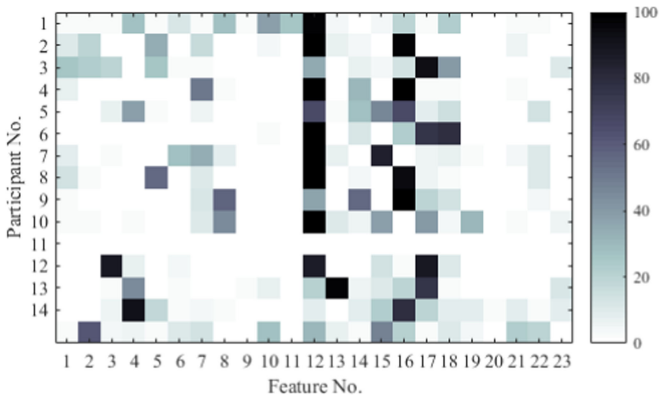


Fig. 7. Frequency each feature was selected in classifying low/positive vs low/negative states.

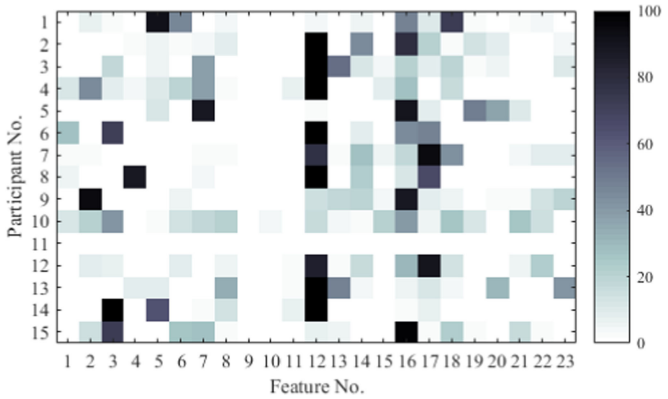


Fig. 8. Frequency each feature was selected in classifying high/positive vs high/negative states.

### 5.2 SAM Assessment

In general, there is a weak agreement between classification accuracies and the accuracy of participant ratings of the images (Fig. 5). These results are in accordance with known emotional processing difficulties in individuals with ASD [6]. From Fig. 5, it is evident that each method of assessment (self *vs* physiological) will vary in effectiveness for each individual. While some participants were able to accurately identify the emotions elicited according to the labels of the image databases, physiological classification proved more effective for others. Certainly, the value of employing an affective-sensitive technology will depend on each individual, and subject-specific methods will be necessary. The varied relationship between SAM accuracies and classification accuracies may also be attributed to variability in stimulus potency. From Figs. 5 and 6, it is evident that participant's success in rating the images depended on the valence/arousal level of the image. Highly negative images were excluded in picture selection, as they were unsuitable for the age range considered in this work. A tendency to misinterpret negative stimuli in ASD aligns with previous findings [53]. While positive images were also subject to such constraints, suitable alternatives to potentially offensive or inappropriate images were more readily available. Factors such as attention levels, fatigue, and different methods of emotional inductions may also play a role in the results. Further study is necessary to identify which factors play infleuenctial roles in our ability to classify emotional states on a physiological-level and assessing which individuals would benefit from an emotion sensing technology.

With regards to parent rating, several parents noted that they had difficulty in rating their child's emotional state due to their child's lack of expression throughout the task. Thus, the accuracy of the parents' rating and the agreement between participant and parent ratings is low for the majority of the participants (Fig. 6).

### 5.3 Frequently Selected Features

In this work we, considered a large number of physiological features to classify affective states. Due to our single-trial classification approach, we were able to employ a wide candidate feature pool and through feature selection, select the optimal set of features and train a classifier on a per-subject basis. Selected features can provide insight into the value of each feature type, and in turn, the physiological signal(s)

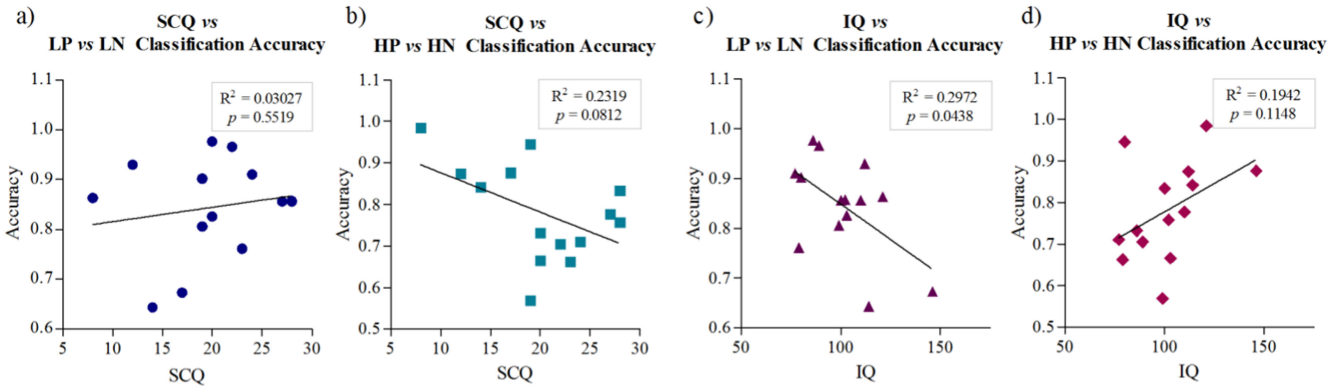


Fig. 9. Relationship between classification accuracy for a) LP vs LN or b) HP vs. HN accuracies and SCQ scores, and c) LP vs LN or d) HP vs. HN accuracies and IQ scores.

relevant to differentiating affective states. The mean SC was most frequently selected across all participants. Interestingly, SC is solely controlled by the sympathetic brain of the ANS, whereas the other signals measured hereon are under both sympathetic and parasympathetic control [51]. Perhaps SC is controlled with higher specificity, facilitating a better reflection of affective responses.

A variety of feature combinations from the four different sensors were selected across participants (Figs. 7 and 8) suggesting each of the physiological measures considered did indeed contain unique, complimentary information. Furthermore, the selected features indicate the optimal physiological signals are subject-dependent. This aligns with literature pertaining to person stereotypy in ASD, where autonomic changes can manifest as unique, but reproducible changes in various physiological measurements [52]. This variability in optimal features highlights the importance of developing subject-specific technologies and individual per-subject analyses when considering ASD.

Variations in the features selected across participants may also be due, in part, to differences in signal quality. Position and adhesion of the sensors, and the degree of noise in the signals can vary across participants, and potentially affect the classifiable information contained in each measure. Our findings suggest that classification of emotional states should accommodate individuals' differences in autonomic reactivity through subject-specific classification, rather than solely focusing on group-averaged changes.

### 5.4 Relations Between Demgraphic Variables/ Behavioural Measures and Accuracy

Fig. 9 shows the relationships between classification accuracies and SCQ and IQ scores. The only significant relationship was found between IQ and classification accuracy of low positive/low negative states ($r^2 = 0.2972$, $p = 0.0438$). This moderate negative linear relationship indicates that a higher IQ can challenge our ability to classify states of emotional valence during low arousal states. This may due to a lower overall response during low intensity states with increasing levels of intelligence, or higher variability among response observed during low levels of arousal. A higher SCG indicates higher severity of ASD symptomatology.

The data suggest that ASD symptomatology does not strongly affect the ability automatically recognize classify affective states using physiological measures.

### 5.5 Comparison to Previous Work

To the best of our knowledge, the present study is the first to explore automatic physiologically-based classification along dimensions of arousal and valence in individuals with ASD. While the works of Groden et al. [37], and Ben-Shalom et al. [11] considered arousal in children with ASD using physiological measures, arousal levels were compared on a group-averaged, rather than single-trial basis. Similar to our work, Liu et al. [38] and Kushki et al. [36] explored automatic discrimination of arousal levels on a single-trial basis. Liu et al. [38] reported average accuracies similar to our work, while Kushki et al. [36] reported average accuracies higher than we achieved (95 percent). However, both these studies considered classification problems different from ours. Liu et al. considered differentiation of specific affective states of liking, anxiety and engagement induced by two computer-based cognitive tasks. Game parameters were manipulated to induce various levels of the affective states. Similarly, Kushki et al. considered the identification of anxiety-related arousal induced by two anxiogenic tasks that required specific performance (Stroop task and public speaking). These overt tasks likely induced a more consistent response across participants relative to the image-viewing task considered in our work. The more subjective nature of our task may, in part, account for differences in accuracies between our work and that of [38] and [36].

## 6 Conclusions, Limitations and Future Work

In this study, we investigated the feasibility of differentiating physiological activity collected while viewing images known to elicit affective states along dimensions of valence (positive, negative) and arousal (high, low) in children with ASD using physiological signals. We were able to classify measurements collected while viewing images of a high/positive state from a high/negative state at an average accuracy of 78.1 percent, and a low/positive state from a low/negative state at an average accuracy of 84.5 percent. Our findings suggest there are indeed discernable patterns within the autonomic measures considered that can be classified on a single-trial basis to potentially identify different affective states. Furthermore, our results suggest that both linear and non-linear classifiers can be of value in addressing classification problems of affective states.

The findings of this study indicate the promise of establishing a physiologically-based approach to emotion recognition for children with ASD that is non-invasive, economically-feasible and language free. Such a system could be used to help individuals improve individuals' awareness of their emotions through feedback reflecting their body state. Additionally, these findings and continued research in this area will improve our understanding of emotional processing challenges among individuals with ASD.

This study has the following limitations. Firstly, we were ethically restricted in the stimuli we were able to use. Exploring a greater number of stimuli would enhance the robustness of the classifier and enhance generalizability to natural social settings consisting of a wide degree of potential stimuli with varying degrees of potency. Second, data collection was limited in duration, restricting the information that could be extracted from the physiological signals (*e.g.*, frequency-domain features from ECG signals were not considered). The duration of the sessions was limited to minimize potential variations in attention to stimuli and adherence to the protocol across the session. Although attention to the stimuli was confirmed by experimenter observation for all participants, variations in attention with and across participants may contribute, in part, to variability in results. The effect of factors such as fatigue and concentration on classification accuracies could be better elucidated by within-subject comparisons in a multi-session study. Furthermore, habituation to repeated exposure to emotional stimuli could potentially diminish the elicited response within a block, or with the experimental session, and would diminish classification accuracies. Further investigation of this factor ought to be explored in longer-term studies. Furthermore, our sample size was limited. This impeded, for example, our ability to consider subgroups of participants. Our results, including variability in optimal features selected across participants and the need for both linear and non-linear classifiers to optimize classification, suggest a need for individual-specific approaches in detecting and classifying affective states in ASD. Further investigation would indicate how this variability would affect the potential of establishing a global classifier to establish an assistive technology suitable to use across the population, or the paradigm necessary to develop individualized technologies. While a larger number of participants would better indicate the generalizability of the classification results presented, the accuracies obtained demonstrate the feasibility of classifying affective states in ASD in a single-trial basis. Future work ought to consider a larger sample size in order identify homogeneoussubgroups of individuals with the use of clustering. Additionally, classification accuracies achieved in this work represent classification of autonomic responsivity to positive and negative emotional stimuli in ASD only. Given the difficulties in self-evaluation of emotions, a key challenge in studying emotional responses in children with ASD is evaluating the true emotional states induced. There is currently no consensus in the literature as to how this challenge is best addressed. A comparison of physiological responses to affective stimuli in ASD and typically develo ping controls could heighten our understanding of ASD. However, establishing a ground-truth in physiological reactivity for the stimuli and paradigm employed in this work was not the focus our focus. Rather, our aim was evaluating an objective means of automatically detecting emotional responses to affective stimuli for eventual use in therapeutic technologies for ASD. Given that ASD is associated with differences in emotional experiences and autonomic responses [56], [57], there is no guarantee participants with ASD will experience the stimuli in the same way as typically developing populations, and in turn, the classification paradigm would be the same. Furthermore, the stimuli used in this study have been validated with neurotypical populations. The International Affective Picture System (IAPS) and the Geneva Affective Picture Database (GAPED) each contain a standardized set of images with normative ratings of valence and arousal, specifically design and validated for experimental investigations of emotions. Future work ought to consider a comparison to classification accuracies achievable with typically developing children to determine whether physiological patterns can be better classified in one group than the other, and the potential of establishing a global classifier of emotional states. Lastly, within-subject variability across multiple sessions ought to be considered in a longer-term study. A multi-session study would also permit collection of a larger number samples and therefore, consideration of higher order classification problems (*i.e.*, ternary and quaternary problems).

## Acknowledgments

This study was funded by the Kids Brain Health Network and The Networks of Centres of Excellence program.

## References

[1] American Psychiatric Association, *Diagnostic and statistical manual of mental disorders: DSM-5$^{TM}$*, Fifth edition. Arlington, VA: American Psychiatric Association, 2013.

[2] K. Honkalampi, J. Hintikka, A. Tanskanen, J. Lehtonen, and H. Viinamäki, "Depression is strongly associated with alexithymia in the general population," *J. Psychosom. Res.*, vol. 48, no. 1, pp. 99–104, 2000.

[3] S. Baron-Cohen, "Understanding other minds: Perspectives from autism," *Neurology*, vol. 45, no. 1, pp. 210–210, 1995.

[4] U. Frith, "Emanuel Miller lecture: Confusions and controversies about Asperger syndrome," *J. Child Psychology Psychiatry Allied Disciplines*, vol. 45, no. 4. pp. 672–686, 2004.

[5] S. Baron-Cohen, M. V. Lomardo, and H. Tager-Flusberg, *Understanding Other Minds: Perspectives Develop. Social Neuroscience*, 3rd Edition. New York, NY: Oxford Scholarship Online, 2013.

[6] G. Silani, G. Bird, R. Brindley, T. Singer, C. Frith, and U. Frith, "Levels of emotional awareness and autism: An fMRI study," *Soc. Neurosci.*, vol. 3, no. 2, pp. 97–112, 2008.

[7] P. E. Sifneos, "The prevalence of 'alexithymic'characteristics in psychosomatic patients," *Psychother. Psychosom.*, vol. 22, pp. 255–262, 1973.

[8] W. Linden, F. Wen, and D. L. Paulhus, "Measuring alexithymia: Reliability, validity, and prevalence," in *Advances Personality Assessment*, Hillsdale, NJ.: Erlbaum, vol. 10, pp. 51–95, 1995.

[9] E. Hill, S. Berthoz, and U. Frith, "Brief report: Cognitive processing of own emotions in individuals with autistic spectrum disorder and in their relatives brief report: Cognitive processing of own emotions in individuals with autistic spectrum disorder and in their relatives," *J. Autism Develop. Disorders*, vol. 34, no. 2, pp. 229–235, 2004.

[10] J. A. Lambie and A. J. Marcel, "Consciousness and the varieties of emotion experience: A theoretical framework," *Psychol. Rev.*, vol. 109, no. 2, pp. 219–259, 2002.

[11] D. Ben Shalom, et al., "Normal physiological emotions but differences in expression of conscious feelings in children with high-functioning autism," *J. Autism Dev. Disord.*, vol. 36, no. 3, pp. 395–400, 2006.

[12] D. P. Kennedy, E. Redcay, and E. Courchesne, "Failing to deactivate: resting functional abnormalities in autism," *Proc. Nat. Acad. Sci. United States America*, vol. 103, no. 21, pp. 8275–8280, 2006.

[13] C. Rieffe, M. Meerum Terwogt, and K. Kotronopoulou, "Awareness of single and multiple emotions in high-functioning children with autism," *J. Autism Dev. Disord.*, vol. 37, no. 3, pp. 455–465, 2007.
[14] J. L. Andreassi, *Psychophysiology: Human Behavior & Physiological Response*. Hove, United Kingdom: Psychology Press, 2013.
[15] S. Luu and T. Chau, "Decoding subjective preference from single-trial near-infrared spectroscopy signals," *J. Neural Eng.*, vol. 6, no. 1, 2009, Art. no. 016003.
[16] E. R. Kandel, J. H. Schwartz, and T. M. Jessell, *Principles of Neural Science*. New York: NY, USA: McGraw-Hill, vol. 3, pp. 1425–1400, 2000.
[17] F. Nasoz, C. L. Lisetti, K. Alvarez, and N. Finkelstein, "Emotion recognition from physiological signals for user modeling of affect," in *Proc. User Model*, 2003, Art. no. 8.
[18] P. Ekman, et al., "Universals and cultural differences in the judgments of facial expressions of emotion," *J. Pers. Soc. Psychol.*, vol. 53, no. 4, pp. 712–717, 1987.
[19] K. H. Kim, S. W. Bang, and S. R. Kim, "Emotion recognition system using short-term monitoring of physiological signals," *Med. Biol. Eng. Comput.*, vol. 42, no. 3, pp. 419–427, 2004.
[20] J. Kim and E. André, "Emotion recognition based on physiological changes in music listening," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2067–2083, Dec. 2008.
[21] D. Kulic and E. A. Croft, "Affective state estimation for human–robot interaction," *IEEE Trans. Robot.*, vol. 23, no. 5, pp. 991–1000, Oct. 2007.
[22] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 10, pp. 1175–1191, Oct. 2001.
[23] G. Rigas, C. D. Katsis, G. Ganiatsas, and D. I. Fotiadis, "A user independent, biosignal based, emotion recognition method," *User Model 2007*, vol. 4511, pp. 314–318, 2007.
[24] A. Haag, S. Goronzy, P. Schaich, and J. Williams, "Emotion recognition using bio-sensors: First steps towards an automatic system," *Affect. Dialogue Syst.*, vol. i, pp. 36–48, 2004.
[25] F. Nasoz, K. Alvarez, C. Lisetti, and N. Finkelstein, "Emotion recognition from physiological signals using wireless sensors for presence technologies," *Cognition Technol. Work*, vol. 6, no. 1, pp. 4–14, 2004.
[26] C. Zong and M. Chetouani, "Hilbert-Huang transform based physiological signals analysis for emotion recognition," *Proc. IEEE Int. Symp. Signal Process. Inf. Technol. ISSPIT*, 2009, pp. 334–339.
[27] B. Cheng and G. Liu, "Emotion recognition from surface EMG signal using wavelet transform and neural network," in *Proc. 2nd Int. Conf. Bioinf. Biomed. Eng.*, no. 1, 2008, pp. 1363–1366.
[28] F. Hönig, J. Wagner, A. Batliner, and E. Nöth, "Classification of user states with physiological signals: On-line generic features vs. specialized feature sets," in *Proc. Eur. Signal Process. Conf.*, 2009, pp. 2357–2361.
[29] F. Agrafioti, D. Hatzinakos, S. Member, and A. K. Anderson, "ECG pattern analysis for emotion detection.pdf," *IEEE Trans. Affective Comput.* vol. 3, no. 1, pp. 102–115, Jan.-Mar. 2012.
[30] L. Shen, M. Wang, and R. Shen, "Affective e-learning: Using 'emotional' data to improve learning in pervasive learning environment related work and the pervasive e-learning platform," *Educ. Technol. Soc.*, vol. 12, pp. 176–189, 2009.
[31] G. Chanel, C. Rebetez, M. Bétrancourt, and T. Pun, "Emotion assessment from physiological signals for adaptation of game difficulty," *IEEE Trans. Syst. Man Cybern. Part A: Syst. Hum.*, vol. 41, no. 6, pp. 1052–1063, Nov. 2011.
[32] J. Zhai and A. Barreto, "Stress detection in computer users based on digital signal processing of noninvasive physiological variables," in *Proc. 28th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, no. Ic, 2006, pp. 1355–1358.
[33] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis using physiological signals," *IEEE Trans. Affective Comput.* vol. 3, no. 1, pp. 18–31, Jan.-Mar. 2012.
[34] J. Scheirer, R. Fernandez, J. Klein, and R. W. Picard, "Frustrating the user on purpose: A step toward building an affective computer," *Interact. Comput.*, vol. 14, no. 2, pp. 93–118, 2002.
[35] K. H. Kim, S. W. Bang, and S. R. Kim, "Emotion recognition system using short term monitoring of physiological signals," *Med. Biol. Eng. Comput.*, vol. 42, no. 3, pp. 419–427, 2004.
[36] A. Kushki, A. Khan, J. Brian, and E. Anagnostou, "A Kalman filtering framework for physiological detection of anxiety in children with autism spectrum disorder," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 3, pp. 990–1000, Dec. 2014.
[37] J. Groden, et al., "Assessing cardiovascular responses to stressors in individuals with autism spectrum disorders," *Focus Autism Dev. Disabil.*, vol. 20, no. 4, pp. 244–252, 2005.
[38] C. Liu, K. Conn, N. Sarkar, and W. Stone, "Physiology-based affect recognition for computer-assisted intervention of children with autism spectrum disorder," *Int. J. Hum.-Comput. Stud.*, vol. 66, no. 9, pp. 662–677, 2008.
[39] C. Lord et al., "The autism diagnostic schedule – generic: A standard measures of social and communication deficits associated with the spectrum of autism," *J. Autism Dev. Disord.*, vol. 30, no. 3, pp. 205–223, 2000.
[40] C. Maaoui and A. Pruski, "Emotion recognition through physiological signals for human-machine communication," *Cut. Edge Robot.*, pp. 317–333, 2010.
[41] Y. Gu, S.-L. Tan, K.-J. Wong, M.-H. R. Ho, and L. Qu, "A biometric signature based system for improved emotion recognition using physiological responses from multiple subjects," in *Proc. 8th IEEE Int. Conf. Ind. Inf.*, 2010, pp. 61–66.
[42] E. S. Dan-Glauser and K. R. Scherer, "The Geneva affective picture database (GAPED): A new 730-picture database focusing on valence and normative significance," *Behav. Res. Methods*, vol. 43, no. 2, pp. 468–77, 2011.
[43] S. Bölte, S. Feineis-Matthews, and F. Poustka, "Brief report: Emotional processing in high-functioning autism - physiological reactivity and affective report," *J. Autism Dev. Disord.*, vol. 38, no. 4, pp. 776–781, 2008.
[44] N. Jatupaiboon, S. Pan-Ngum, and P. Israsena, "Real-time EEG-based happiness detection system," *Sci. World J.*, vol. 2013, 2013, Art. no. 618649.
[45] J. Pan and W. J. Tompkins, "A real-time qrs detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985.
[46] [Online]. Available: http://biosig.sourceforge.net/download.html
[47] W. F. Ganong, *Review of Medical Physiology*. 22nd Edition. McGraw-Hill Companies, New York, p. 928, 2005.
[48] P. Pudil, J. Novovicova, and J. Kittler, "Floating search methods in feature selection," *Pattern Recog. Lett.*, vol. 15, no. 11, pp. 1119–1125, Nov. 1994.
[49] G. Muller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller, "Better than random? A closer look on BCI results," *Int. J. Bioelectromagn*, vol. 10, no. 1, pp. 52–55, 2008.
[50] C. Ashwin, E. Chapman, L. Colle, and S. Baron-Cohen, "Impaired recognition of negative basic emotions in autism: A test of the amygdala theory," *Soc. Neurosci.*, vol. 1, no. 3–4, pp. 349–363, 2006.
[51] W. Boucsein, *Electrodermal Activity*. New York, NY, USA: Springer, vol. 3. 2012, Art. no. 618.
[52] J. I. Lacey and B. C. Lacey, "Verification and extension of the principle of autonomic response-stereotypy," *Am. J. Psychol.*, vol. 71, no. 1, pp. 50–73, 1958.

**Sarah Sarabadani** received the master's of applied science degree in biomedical engineering from the University of Toronto, Toronto, Ontario, Canada, in 2016. She studied mechanical engineering with the Sharif University of Technology, Tehran, Iran.

**Larissa C. Schudlo** received the bachelor's of engineering degree in electrical and biomedical engineering from McMaster University, Hamilton, ON, Canada, in 2010, and the master's of applied science and doctorate degrees in biomedical engineering from the University of Toronto, Toronto, ON, Canada, in 2012 and 2016, respectively. Her research focuses on classification of physiological and functional near-infrared spesctroscopy signals in neuro atypical popoulations. She is currently an assistant professor with Ryerson University, Toronto, ON, Canada.

**Ali Akbar Samadani** received the bachelor's of science degree from the University of Kuwait, in 2006 and the master's of science degree in electrical and computer engineering from the University of Manitoba, in 2009, and the doctorate degree in electrical and computer engineering from the University of Waterloo, in 2014. He was an NSERC post-doctoral fellow with the University of Toronto and Holland Bloorview Kids Rehabilitation Hospital, Toronto, ON, Canada from 2014-2016. He is currently employed with Philips, Boston, Massachusetts. His research interests lie in machine learning and biological signal processing. He is a member of the IEEE.

**Azadeh Kushki** received the bachelor's of applied science degree in computer engineering, and the master's of applied science and doctorate degrees in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada. She completed post-doctoral work with McMaster University, Hamilton, ON, Canada and the University of Toronto. She is currently a scientist with Holland Bloorview Kids Rehabilitation Hospital, Toronto, ON, Canada and an assistant professor with the University of Toronto. Her research focuses on classification of physiological signals and technology development for children with Autism Spectrum Disorder. She is a member of the IEEE.