Behavioral and physiological differences during an emotion-evoking task in children at increased likelihood for autism spectrum disorder

Lori-Ann R. Sacrey, Lonnie Zwaigenbaum, Jessica A. Brian, Isabel M. Smith, Vickie Armstrong, Tracy Vaillancourt and Louis A. Schmidt

Abstract

Literature examining emotional regulation in infants with autism spectrum disorder (ASD) has focused on parent report. We examined behavioral and physiological responses during an emotion-evoking task designed to elicit emotional states in infants. Infants at an increased likelihood for ASD (IL; have an older sibling with ASD; 96 not classified; 29 classified with ASD at age two) and low likelihood (LL; no family history of ASD; n = 61) completed the task at 6, 12, and 18 months. The main findings were (1) the IL-ASD group displayed higher levels of negative affect during toy removal and negative tasks compared to the IL non-ASD and LL groups, respectively, (2) the IL-ASD group spent more time looking at the baseline task compared to the other two groups, and (3) the IL-ASD group showed a greater increase in heart rate from baseline during the toy removal and negative tasks compared to the LL group. These results suggest that IL children who are classified as ASD at 24 months show differences in affect, gaze, and heart rate during an emotion-evoking task, with potential implications for understanding mechanisms related to emerging ASD.

Keywords: affect; ASD; autism; baby sibling; gaze; heart rate; physiology

(Received 4 May 2022; revised 19 October 2022; accepted 20 October 2022)

Introduction

Emotional regulation (ER) is the ability to manage the intensity and valence of our emotional reactions to internal and external stimuli (Cole et al., 1994; Thompson, 1994) and is associated with the development of social skills (Eisenberg et al., 2000, 2010), the onset of behavioral problems (Nolan et al., 2001; Upshur et al., 2009), as well as academic ability (Blair & Razza, 2007; Welsh et al., 2010). Much of the literature on ER in infants and toddlers has relied on parent reports (Mazefsky et al., 2013, 2021), which may limit our ability to investigate age-related changes due to a lack of standardization of contexts and responding based on social desirability (Achenbach, 2011; De Los Reyes et al., 2015). There are a growing number of published reports that include direct observation of children during tasks that evoke affective responses, with facial affect (Busuito et al., 2019; Buss et al., 2005; Ham & Tronick, 2009), gaze (Mireault et al., 2018; Sacrey et al., 2021b), and heart rate (Fox et al., 2000; Propper & Moore, 2006) serving as markers of ER. A review of physiological measurement during emotionally salient tasks in neurotypical children between 4 and 48 months of age found that resting heart rate decreases with age, heart rate increases during negatively valanced emotion-evoking (EE) tasks (when compared to heart rate at rest), and heart rate is associated with measures of facial affect and gaze (Sacrey et al., 2021a). Thus, elucidating early differences in ER may help inform our understanding of normative and atypical trajectories of child development.

Children with autism spectrum disorder (ASD) and their younger siblings, who are at increased likelihood (IL) of also being diagnosed with ASD, show differences in ER relative to similar-aged peers. That is, they display higher rates of negative emotions, such as sadness and fear, and lower rates of positive emotions on parent-reported questionnaires (Ben Shalom et al., 2006; Capps et al., 1993; Garon et al., 2016; Hirschler-Guttenberg et al., 2015; Putnam et al., 2006; Samson, 2013). Direct observations of ER in IL siblings show a similar pattern of findings; they display greater levels of displayed fear (Macari et al., 2018) and negative affect (Sacrey et al., 2021b) during emotion-elliciting tasks, and lower rates of positive affect during free play (Filliter et al., 2015) compared to neurotypical peers. An under-explored area of IL in siblings show a similar pattern of findings; they display greater levels of displayed fear (Macari et al., 2018) and negative affect (Sacrey et al., 2021b) during emotion-elliciting tasks, and lower rates of positive affect during free play (Filliter et al., 2015) compared to neurotypical peers. An under-explored area of IL in siblings is physiological responses to EE tasks (e.g., toy removal following toy play; Goldsmith & Rothbart, 1991). Atypical arousal (e.g., differences in heart rate reactivity compared to neurotypical peers) is well documented in older children and adults with ASD (Lydon et al., 2016). Because physiological indices of arousal can register differences in early-developing processes, they may be informative for mechanisms underlying ER,
development and ASD symptom emergence (Feldman, 2009; Kushki et al., 2013; Neuhaus et al., 2014).

In the present study, we examined behavioral and physiological responses to EE tasks (bubbles, toy play, toy removal, masks, face washing, and hair brushing) at 6, 12, and 18 months of age in children who were at low likelihood (LL; no family history of ASD) and IL for ASD. Children participated in an EE task while wearing heart rate sensors and video recordings were coded for facial affect and on-task gaze. All participants underwent an assessment for ASD at 24 months, then categorized into three groups: LL, IL not classified with ASD (IL non-ASD), and IL classified with ASD (IL-ASD). We predicted that, when compared to the LL and IL non-ASD groups, the IL-ASD group would (1) show greater increases in heart rate from baseline, (2) display higher levels of negative affect and lower levels of positive affect; and (3) spend less time looking at on-task objects during the EE Task.

Method

Participants

Infant siblings of children with ASD were recruited between the ages of 6 and 12 months from families attending one of three multidisciplinary ASD clinical centers and surrounding communities [Glenrose Rehabilitation Hospital (Edmonton, AB), Holland Bloorview Children’s Hospital (Toronto, ON), and IWK Health Centre (Halifax, NS)]. Participants were assessed at 6, 12, 18, and 24 months of age. The research ethics board at each institution approved this study and all families gave written informed consent prior to study enrollment.

For the IL group, diagnosis of ASD in the older sibling (i.e., proband) was confirmed by a review of diagnostic records using DSM-5 criteria (APA, 2013). No IL infant had any identifiable neurological or genetic conditions, nor severe sensory or motor impairments. Infants at LL were recruited from the same communities and had at least one older sibling but no reported first- or second-degree relatives with an ASD diagnosis. All participants were born between 36 and 42 weeks of gestation, with birth weights greater than 2,500 g.

EE task

Positive and negative affect, as well as gaze, were measured using tasks adapted from the Laboratory Temperament Assessment Battery (Lab-TAB; Goldsmith & Rothbart, 1991), a comprehensive temperament assessment that includes episodes designed to elicit behavior related to differing dimensions of temperament, including smiling, reaching, crying, touching, or changes in facial expression. The EE Task (Sacrey et al., 2021b) was completed at 6, 12, and 18 months of age. All EE Task data were collected prior to the COVID pandemic.

EE task setup

Children were seated in a high-chair at a height-adjustable table with their parent seated to their right, according to the parent location guidelines for the Mask and Toy Removal tasks in the Lab-TAB manual (Goldsmith & Rothbart, 1991). All phases of the EE Task, including the Baseline video, occurred with the child seated in the high-chair. The Baseline videos were shown on a laptop or computer monitor placed on the table in front of the child (see Figure 1). Once the video ended, the computer/monitor was removed to the floor out of the child’s sight. The objects used for each task were held in an opaque bin beside the examiner, also out of the child’s sight. The phases included within our EE Task are shown in Figure 1:

1. Baseline 1 phase – The child was shown a 2-min video comprising 15-s clips of intermixed screensaver images and “Baby Einstein” clips accompanied by instrumental music to allow an opportunity to acclimate to the research setting (neutral task). Examiners faced neutral throughout the video playback.
2. Bubbles phase – The experimenter blew bubbles toward the child and directed the child’s attention toward the bubbles for 90 s (positive task). Examiners faced positive (smiling) throughout the phase.
3. Baseline 2 phase – The child was shown the same 2-min video from Baseline 1 to allow an opportunity to return to baseline (neutral task). Examiners faced neutral throughout the video playback.
4. Toy Play phase – The child was given a toy for 30 s that lit up and made music when buttons were pushed (positive task). Examiners faced neutral throughout the phase.
5. Toy Removal phase – An appealing toy (used in the Toy Play phase) was moved out of the child’s reach but within sight for 30 s (negative task). Examiners faced neutral throughout the phase.
6. Negative Tasks phase – This phase comprised three sub-tasks: (1) The experimenter wore a blank mask on their face and sat still and quietly for 15 s before switching the blank mask for a cow mask and sitting for 15 more seconds (Masks). (2) The experimenter brushed the child’s hair with a comb or soft brush for 15 s (Hairbrush). (3) The experimenter gently wiped the child’s face (forehead, cheeks, chin, nose) with a baby wipe for 15 s (Face wash). Examiners faced neutral throughout the phase.
7. Baseline 3 phase – The child watched the same 2-min video from Baselines 1 and 2 to allow an opportunity to return to baseline following the negative tasks. Examiners faced neutral throughout the video playback.

Affect and gaze coding

The EE Task was video-recorded and affect and gaze were coded offline from video recordings using Noldus Observer 13 XT behavioral coding software (see Table S1 in Supplemental materials for brief coding scheme). Coding was completed in two separate viewings of the entire video recording for each participant. The first viewing involved marking onset and offset of each task phase as well as coding for facial affect and in the second viewing gaze was coded. Videos were played at real time for coding purposes. Phases were coded continuously and codes were mutually exclusive and exhaustive such that one code ended the previous code. Periods between phases were coded as “transition” episodes and were not coded for behavior or included in any analyses.

Affect. Affect was coded in 5-s intervals as either negative, neutral, or positive on a 5-point scale from −2 to +2 based on both facial and vocal cues. Periods during which the face was not visible and vocal cues for affect were absent were coded as “not codable” (for definitions associated with use of facial or vocal cues alone to code affect, see Supplemental materials). Interval coding was selected because the onset and offset of affect intensity was difficult to define, as facial affect cues can change rapidly. Mean affect was calculated for each phase of the EE Task by taking the mean of all 5-s intervals. For example, the Toy Play phase was 30 s and comprised 6 coded 5-s intervals. The mean affect for the Toy Play phase was calculated as the sum of the codes for each of the 6 intervals divided by 6.
Gaze. Gaze was coded continuously with codes being mutually exclusive and exhaustive. The behavior of interest was the target of the child’s gaze. This included looking at “on-task” and “off-task” objects, the experimenter conducting the task, the parent sitting beside the child, and any gaze aversion. On-task gaze objects included the computer monitor for the Baseline phases, bubbles or bubble wand for the Bubbles phase, the toy used for the Toy Play and Toy Removal phases (same toy), and the two masks, the comb/brush, and the wipe used in the sub-phases of the Negative Tasks. Off-task objects included nearby objects that the infant manipulated or interacted with (e.g., sensors and cables, as well as objects that parents may have given their children unexpectedly, such as toys or sippy cups, which were removed as quickly as possible). “Other” was used to code any other looking behavior (e.g., scanning the room). We only assessed the on-task gaze behavior. The variable for percentage of time spent on the on-task object was calculated for each phase of the EE Task using the following formula (Sacrey et al., 2021b):

\[
\text{Percentage of time} = \frac{\text{time spent looking at on-task object}}{\text{length of phase}} \times 100
\]

Inter-rater reliability
Two raters coded 20% of the videos to assess reliability using Cohen’s kappa (κ), with 0.01–0.20 representing no to slight agreement, 0.21–0.4 representing fair agreement, 0.41–0.60 as moderate agreement, 0.61–0.80 representing substantial agreement, and 0.81–1.00 representing almost perfect agreement (Marston, 2010). When reliability was assessed using a modifier margin of 1 (codes were within ±1 point between raters), \( \kappa = 95\% \). For gaze, \( \kappa = 89\% \) was achieved when calculating the percentage agreement for duration of gaze codes for the two raters. Both raters were blinded to enrollment group (IL vs LL) and ASD symptom history, but one rater was involved in study visits at one site.

Physiological (electrocardiogram [ECG]) arousal
Three ECG sensors were attached to the child in an inverted triangle, with the right lead placed under the right clavicle, the left lead placed under the left clavicle (both at mid-clavicular line within the rib cage frame), and the ground lead at the lower left abdomen within the rib cage frame. Physiological data were acquired using a ProComp Infinity Encoder (T7500M) and Biograph Infinity Software (Version 6) and sampled at 2048 Hz. The ECG time series that was demarked by task onset and offsets (described in the EE Task setup) was processed as follows. First, the time series was visually inspected for quality (records with greater than 5% noise failed quality control). Next, RR intervals were extracted from the ECG time series using an adapted version of the Pan-Tompkins algorithm (Pan & Tompkins, 1985; Hamilton & Tompkins, 1986) and values outside of the 1.5*interquartile envelope were removed. Finally, heart rate was computed as the inverse of the RR series (beats per minute [bpm]). Heart rate reactivity was calculated for each task by subtracting mean heart rate from baseline.
rate during Baseline 1 from mean heart rate during Bubbles (positively salient), Toy Play (positively salient), Toy Removal (negatively salient), and Negative tasks (negatively salient).

To synchronize the video and ECG record, we recorded and digitized a second channel containing a synchronization signal from Noldus Observer 13.0 (Sync Channel). The synchronized channel contained an on and off pulse that occurred when the video recording started and stopped. The signal was sent from Observer using the computer’s COM port to a voltage isolator. The voltage isolator, in turn, sent the on-off signal to the ProComp Infinity Encoder. Following processing of the physiology, heart rate and sync signals were imported into Observer to be synchronized with coded behavior.

Physiological data loss due to movement artifacts, removal of heart rate sensors, or discontinuation of EE Task due to increased levels of negative arousal was 25% at 6 months, 20% at 12 months, and 34% at 18 months. A comparison of data loss by group showed no differences among the three groups at 6 (X²(2) = 1.46, p = .48), 12 (X²(2) = 2.03, p = .36), or 18 months (X²(2) = 1.99, p = .37).

24-month clinical assessment

Children in the IL and LL groups were assessed for ASD diagnosis based on parent report, developmental skills, and ASD symptoms. In-person visits included the Mullen Scales of Early Learning (Mullen, 1995) and the toddler module of the Autism Diagnostic Observation Schedule – 2nd Edition (ADOS-2; Lord et al., 2012; n = 131). A subset of children was unable to attend in-person visits due to COVID-19 restrictions and were administered the TELE-ASD-PEDS (Corona et al., 2021; n = 55). Parents who participated in either in-person and virtual assessment completed the Parent Concern Form (Sacrey et al., 2015) and the Vineland Adaptive Behavior Scales III (Sparrow et al., 2016). We compared in-person and virtual completers on all independent variables at 6, 12, and 18 months of age.

Mullen scales of early learning (Mullen)

The Mullen (Mullen, 1995) is a directly administered developmental measure that assesses Visual Reception, Receptive Language, Expressive Language, Fine Motor, and Gross Motor abilities; an Early Learning Composite comprises the first four scales. We administered the Mullen at 24 months.

Autism Diagnostic Observation Schedule – 2nd edition (ADOS-2)

The ADOS-2 (Lord et al., 2012) was administered by a research-reliable examiner; it includes standardized activities and “presses” intended to elicit communication, social interaction, imaginative use of play materials, and repetitive behavior. The Toddler module was administered at the 24-month assessment, and Social Affect (SA), Restricted and Repetitive Behavior (RRB), and Total algorithm scores were derived.

Parent concerns form

This is a semi-structured interview that collects information about parent concerns related to ASD in the first 2 years (Sacrey et al., 2015). At each timepoint, parents were asked if they had current concerns in each of three broad areas: (1) general (sleep, diet, sensory, motor), (2) behavioral (social, play, behavioral problems, repetitive behaviors/restricted interests), and (3) communication (verbal/nonverbal, regression). The data are not included in this paper but rather were used to inform 24-month classifications.

Vineland Adaptive Behavior Scales, 3rd edition

The Vineland Adaptive Behavior Scales, 3rd edition assesses child adaptive behavior in the communication, socialization, daily living skills, and motor, domains. The Survey Interview (age range: birth to 90 years) is administered to a parent using a semi-structured interview. The Vineland was administered at 24 months of age (Sparrow et al., 2016).

TELE-ASD-PEDS

Participants who were unable to attend an in-person assessment at 24 months due to COVID-19 restrictions participated in a telemedicine-based ASD assessment for toddlers, the TELE-ASD-PEDS (Corona et al., 2021), a virtual assessment for the signs of ASD that is implemented using an online video conferencing software (e.g., Zoom). The assessment presses for socially directed speech and gestures, eye contact, unusual vocalizations or sensory exploration, and repetitive play—all behaviors that help to inform a clinical best estimate of ASD. This assessment was administered at 24 months for virtual participants.

Statistical analysis

All analyses were run in Statistical Package for the Social Sciences (version 24, IBM). The participants were followed longitudinally between 6 and 24 months, thus participants who completed at least one EE Task at 6, 12, or 18 months and had a 24-month assessment were included in data analyses. First, participant demographics were compared between the three groups (IL-ASD, IL non-ASD, LL) using Kruskal–Wallis H tests for continuous variables and chi-square analyses for categorical variables. Second, developmental outcomes (Mullen, Vineland) and ASD signs (ADOS) for participants who completed the 24-month assessment were compared using Kruskal–Wallis H tests. Third, heart rate, affect, and gaze data were standardized using logarithmic transformation (with an added constant of 100 to ensure there were no “0” or negative values) to control for skewness and kurtosis. Fourth, Spearman rho correlations were run between heart rate, gaze, and affect log transformed scores to assess for behavioral and physiological relationships for each EE Task phase (Baseline 1, 2, and 3, Bubbles, Toy Play, Toy Removal, and Negative task phases). Fifth, we ran a series of linear mixed models with group (IL-ASD, IL non-ASD, LL) and age (6, 12, 18) as the independent variables and log transformed scores on each phase of the EE Task (Baseline 1, 2, and 3, Bubbles, Toy Play, Toy Removal, and Negative task phases) for heart rate, affect, and gaze as the dependent variables. The significance level of group and age effects were adjusted using Bonferroni corrections in post hoc analyses.

Participant characteristics

Displayed in Table 1, 61 LL (38 boys and 23 girls), 96 IL non-ASD (45 boys and 51 girls), and 29 IL-ASD (22 boys and 7 girls) children contributed data to this study. No differences were seen between groups based on race/ethnicity, parental marital status, household income, or exact age at 6-, 12-, 18-, or 24-month assessments (all p’s > .05), however there was a sex effect, with a higher proportion of boys to girls in the IL-ASD group compared to the IL non-ASD group (p = .005).
Clinical assessment

Mullen

There were differences between the three groups for Visual Reception (H = 19.88, p < .001), Fine Motor (H = 17.01, p < .001), Receptive Language (H = 18.35, p < .001), and Expressive Language (H = 22.83, p < .001) subscales, as well as the Early Learning Composite (H = 20.65, p < .001). There was no group difference for the Gross Motor subscale (H = 42, p = .81). Post hoc tests revealed that the IL-ASD group differed (had lower scores) from both the IL non-ASD and LL groups for Visual Reception, Fine Motor, Receptive Language, Expressive Language, and the Early Learning Composite (p’s < .001). The IL non-ASD and LL groups did not differ on any subscale or composite (p’s > .05).

Vineland

There were differences between the three groups for Communication (H = 53.22, p < .001), Daily Living (H = 14.90, p < .001), Social (H = 49.11, p < .001), and Motor Skills (H = 12.63, p = .002) subscales, as well as the Adaptive Behavior Composite (H = 55.37, p < .001). Post hoc tests revealed that the IL-ASD group differed (had lower scores) from both the IL non-ASD and LL groups for Visual Reception, Fine Motor, Receptive Language, Expressive Language, and the Early Learning Composite (p’s < .001). The IL non-ASD and LL groups did not differ on any subscale or composite (p’s > .05).

Table 1. Participant characteristics by enrollment group

<table>
<thead>
<tr>
<th>Variable</th>
<th>LL Mean (SD)</th>
<th>IL non-ASD mean (SD)</th>
<th>IL-ASD mean (SD)</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at 6-mo visit, in months</td>
<td>6.53 (.53)</td>
<td>6.41 (.34)</td>
<td>6.79 (1.73)</td>
<td>H = 1.89, p = .39</td>
</tr>
<tr>
<td>Age at 12-mo visit, in months</td>
<td>12.49 (.84)</td>
<td>12.37 (.37)</td>
<td>12.44 (.55)</td>
<td>H = .59, p = .74</td>
</tr>
<tr>
<td>Age at 18-mo visit, in months</td>
<td>18.37 (.39)</td>
<td>18.47 (.37)</td>
<td>18.32 (.19)</td>
<td>H = 4.87, p = .09</td>
</tr>
<tr>
<td>Age at 24-mo visit, in months</td>
<td>24.49 (.49)</td>
<td>24.57 (.64)</td>
<td>24.46 (.42)</td>
<td>H = .80, p = .67</td>
</tr>
<tr>
<td>Sex (n boys:girls)</td>
<td>38:23</td>
<td>45:51</td>
<td>22:7</td>
<td>$^2 = 8.87, p = .01^*</td>
</tr>
<tr>
<td>Child Ethnicity/Race</td>
<td>77% Caucasian</td>
<td>60% Caucasian</td>
<td>54% Caucasian</td>
<td>$^2 = 22.29, p = .22</td>
</tr>
<tr>
<td>Parents’ marital status</td>
<td>77% married</td>
<td>85% married</td>
<td>68% married</td>
<td>$^2 = 7.32, p = .29</td>
</tr>
<tr>
<td>Household income (in Canadian dollars)</td>
<td>5% less than</td>
<td>6% less than</td>
<td>8% less than</td>
<td>$^2 = 27.83, p = .11</td>
</tr>
<tr>
<td></td>
<td>than $40,000</td>
<td>than $40,000</td>
<td>than $40,000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13% $40,001–$80,000</td>
<td>21% $40,001–$80,000</td>
<td>38% $40,001–$80,000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30% $80,001–$125,000</td>
<td>24% $80,001–$125,000</td>
<td>25% $80,001–$125,000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>34% $125,001–$200,000</td>
<td>28% $125,001–$200,000</td>
<td>21% $125,001–$200,000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11% $200,001 and higher</td>
<td>10% $200,001 and higher</td>
<td>4% $200,001 and higher</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7% not given</td>
<td>11% not given</td>
<td>4% not given</td>
<td></td>
</tr>
</tbody>
</table>

*Note: Abbreviations: ASD = autism spectrum disorder; LL = low likelihood; IL = increased likelihood; mo = months; SD = standard deviation.*
Learning Composite (p's < .001). The IL non-ASD and LL groups did not differ on any subscale or composite (p's > .05). The IL-ASD group had lower scores on all subscales and the Adaptive Behaviour Composite compared to the IL non-ASD (p's < .001) and LL (p's < .001) groups, who did not differ (p's < .02).

**ADOS-2**
There was a significant group difference for Total Severity scores (H = 40.13, p < .001), as well as the SA (H = 37.74, p < .001) and the RRB (H = 25.50, p < .001) scores. The IL-ASD group had higher scores on the SA, RRB, and Total Severity scores compared to the IL non-ASD (p's < .001) and LL (p's < .001) groups, who did not differ (p's < .02).

**In-person versus virtual completers**
Affect, on-task gaze, and heart rate variables were compared between the children who completed an in-person versus virtual 24-month assessment using Mann–Whitney U tests with a conservative p value = .01 to account for multiple testing. We found no differences for any heart rate, affect, or gaze measurement between in-person versus virtual group at 6 and 18 months, or for heart rate and affect at 12 months. For on-task gaze, a significant group difference was seen for the Bubbles phase at 12 months (d = .27), with in-person completers spending more time looking at the on-task object (M + SD = 84.34 + 17.24%) compared to virtual completers (M + SD = 80.05 + 11.31%).

**Physiological and behavioral assessment**
Associations between heart rate, affect, and gaze for each phase of the EE Task are presented in Table 2.

**Mean affect**

**Baseline phases**
As shown in Figure 2A, there were no significant effects of group (IL-ASD, IL non-ASD, LL), age (6, 12, 18 months) or group by age interactions for Baseline 1, Baseline 2, or Baseline 3.

**Phases of EE task**

**Group.** A significant effect of group was found for two phases of the EE Task. As shown in Figure 3A, there was a significant effect of group for the Toy Removal phase (F(2,348) = 4.85, p = .008, d = .36), with the IL-ASD group displaying higher levels of negative affect (63% scoring of −1 or −2) compared to the IL non-ASD group (42% scoring a −1 or −2). No other group comparisons were significant. As shown in Figure 4A, there was also a significant effect for the Negative Tasks phase (F(2,335) = 4.42, p = .013, d = .36), with the IL-ASD group displaying higher levels of negative affect (67% scoring −1 or −2) than the LL group (55% scoring −1 or −2). No other group comparisons were significant.

**Age.** There was a significant effect of age for the Bubbles phase (F(2,360) = 20.30, p < .001, d = .57), with higher expressions of positive affect at 18 months (74% scored +1 or +2) compared to 6 (43% scored +1 or +2) and 12 months (62% scored +1 or +2; p's < .001) and higher expressions of positive affect at 12 months compared to 6 months (p = .015).

**Group by age.** There were no significant group by age interactions.

**On-task gaze**

**Baseline phases**

**Group.** As shown in Figure 2B, there was a significant effect for Baseline 2 (F(2,353) = 4.22, p = .015, d = .46), with the IL-ASD group spending more time looking at the screen compared to the IL non-ASD (p = .003) and LL (p = .007) groups, who did not differ (p = .91). There was also a significant effect for Baseline 3 (F(2,333) = 3.90, p = .02, d = .35), with the IL-ASD group spending more time looking at the screen compared to the IL non-ASD group (p = .007). No other group comparisons were significant.

**Age.** A significant effect of age was seen for all baseline phases, with less time spent looking at the screen at 12 months compared to 6 and 18 months. For Baseline 1 (F(2,358) = 8.42, p < .001, d = .18), less time was spent looking at the screen at 12 months compared to 6 (p = .002) and 18 months (p < .001), which did not differ (p = .64). For Baseline 2 (F(2,353) = 7.37, p < .001, d = .26), less time was spent looking at the screen at 12 months compared to 6 (p < .001) and 18 months (p = .002), which did not differ (p = .69). For Baseline 3 (F(2,333) = 4.94, p = .008, d = .18), less time was spent looking at the screen at 12 months compared to 6 (p = .01) and 18 months (p = .006), which did not differ (p = .99).

**Group by age.** There were no significant group by age interactions.

**Phases of EE task**

**Group.** There was a significant effect of group for the Negative Tasks phase (F(2,348) = 4.81, p = .009, d = .33), with the LL group spending more time looking at the task objects compared to the IL-ASD (p = .005) and IL non-ASD (p = .016) groups, who did not differ (p = .26), as shown in Figure 4B.

**Age.** There were no significant effects of Age for any of the EE Task phases.

**Group by age.** There were no significant group by age interactions.

**Heart rate**

**Baseline phases (mean scores)**

**Group.** As shown in Figure 2C, there were no significant group effects.

**Age.** A significant effect of age was seen for Baseline 1 (F(2,240) = 22.23, p < .001, d = .64). Follow-up analyses revealed that heart rate decreased with age, with the highest heart rate at 6 months compared to 12 (p < .001) and 18 months (p < .001), which did not differ (p = .022).

**Group by age.** There were no significant group by age interactions.

**Phases of EE task (reactivity scores)**

**Group.** A significant group effect was seen for two phases of the EE Task. For the Toy Removal phase (F(2,224) = 5.36, p = .005, d = .38), the IL-ASD group showed a greater increase in heart rate from Baseline 1 compared to the LL group (p = .002), as shown in Figure 3C. For the Negative Tasks phase (F(2,223) = 6.85, p < .001, d = .42), the IL-ASD group showed a greater increase in heart rate from Baseline 1 compared to the LL group (p < .001), as shown in Figure 4C.
There was a significant effect of age for the Toy Play phase ($F(2,226) = 4.44$, $p = .013$, $d = .10$), with a greater decrease from Baseline at 18 months compared to 6 months ($p = .005$). No other age comparisons were significant. There was also a significant age effect for the Negative Tasks phase ($F(2,223) = 3.95$, $p = .002$, $d = .30$), with a greater decrease from Baseline at 18 months compared to 6 months ($p = .007$). No other age comparisons were significant.

**Discussion**

We examined behavioral (affect and gaze) and physiological (heart rate) responses during an emotionally salient task at 6, 12, and 18 months in children who were at LL or IL for a later diagnosis of ASD. There were three main results. First, the IL-ASD group displayed higher levels of negative affect during toy removal and negative tasks compared to the IL non-ASD and LL groups, respectively. Second, the IL-ASD group spent more time looking at the baseline task compared to the other two groups. Third, the IL-ASD group showed a greater increase in heart rate from baseline during the toy removal and negative tasks compared to the IL group. Thus, children in the IL-ASD group at 24 months showed differences in affect, gaze, and heart rate responsivity during emotionally salient tasks, compared with the behavior of children who were not classified with ASD.

There were no differences between the three groups for facial affect or gaze during Baseline 1 or positive phases (Bubbles and Toy Play) of the EE Task. When presented with the Toy Removal and Negative Task phases, however, children in the IL-ASD group displayed increased levels of negative affect compared to the other two groups. The IL-ASD group also spent less time looking at the on-task objects during the Negative Task compared to the LL group. Our results are similar to previous research exploring affective reactivity during EE tasks in children at LL and IL for ASD. For example, Macari et al. (2018) also did not find group differences during bubbles and reported increased levels of negative affect in toddlers with ASD during tasks that included masks as a part of the testing protocol (fear protocol). In addition, children with ASD in the Macari et al. (2021) study spent less time looking at distressing stimuli and displayed increased levels of distress compared to their neurotypical peers. Increases in displayed negative affect and reduced gaze during distressing stimuli may reflect atypical regulation and/or expression of negative emotions, potentially affecting early social interactions and contributing to the early presentation of ASD.

We saw no group differences in heart rate during Baseline between the IL and LL groups. This result was similar to previous research that reported no differences in heart rate during a baseline...
video task between preschool children with ASD and neurotypical peers (Bazelmans et al., 2019; Zantinge et al., 2017a, 2017b, 2019). Similarly, our finding that baseline heart rate decreased between 6, 12, and 18 months is consistent with a recent review of physiological responding in neurotypical children. Sacrey et al. (2021a) performed a meta-regression on heart rate during baseline tasks (classified as either watching a video, sitting quietly, a sedentary task, playing with mother, the period immediately before an EE task, or other) and reported successive age-related declines in heart rate between <5 months and 48 months. Baseline heart rate is suggested to reflect an individual’s innate ability to regulate their emotions (Appelhans & Luecken, 2006), suggesting that children who

Figure 2. Affect (A), gaze (B), and heart rate (C) responses during Baseline 1 for the three groups by age.

Figure 3. Affect (A), gaze (B), and heart rate (C) responses during the Toy Removal phase for the three groups by age.
are categorized as IL-ASD show capacity similar to their neurotypical peers. As such, baseline heart rate alone may not be informative as an early biomarker for ASD.

Group differences did emerge during the EE phases of the EE Task, with the IL-ASD group showing a greater increase in heart rate (from baseline) during the Toy Removal and Negative Tasks phases compared to the LL group. These results are similar to those reported by McCormick et al. (2018), who found higher respiratory sinus arrhythmia during a distressing stimulus at 4 months of age in children who were later diagnosed with ASD compared to age-matched controls. Our results contrast with previous reports of no group differences during emotionally salient stimuli (Zantinge et al., 2017a, 2017b, 2019). This inconsistency may be due to methodological differences. For example, Zantinge et al., (2017a, 2019) collected heart rate data from children who were older (41–81 months) than the children in the present study (6–18 months), used different tasks to elicit emotional responses (a robot in Zantinge et al. [2019], a lock box in Zantinge et al. [2017a], and video clip of children arguing in Zantinge et al. [2017b]), and compared actual heart rate values during their EE task rather than change scores (from baseline), as used here. As such, the disparate results may instead stem from age differences. That is, our participants and those of McCormick et al. (2018) were younger than the Zantinge et al. (2017a, 2017b, 2019) samples and may be more reactive because ER systems are still developing (Sacrey et al., 2021a). Moreover, the EE Task used in the present study may be more sensitive to ER differences in children who are IL-ASD.

The EE Task has been previously shown to produce the target emotional responses reliably, with Bubbles and Toy Play producing more positive affect, and Toy Removal and the Negative Tasks producing negative affect (Sacrey et al., 2021b). These results are also supported by the associations between physiological and behavioral measurements during the EE Task, suggesting that emotional responses were measurable and associated on both indices. Similarly, our results also confirm previous results suggesting a relationship between emotional responsivity and ASD symptoms (Sacrey et al., 2021b). On-task gaze and affect during the Toy Removal and Negative Tasks phases at 18 months predicted total ADOS-2 scores at 24 months, whereas a parent-reported measure of ER (Infant Behavior Questionnaire-Revised; IBQ-R) collected at the same timepoint did not (Sacrey et al., 2021b). The results of our study, that children who were in the IL-ASD group did not differ for on measures of heart rate during baseline, but did differ from IL non-ASD and LL groups on two of the negatively valanced emotional tasks are consistent with previous literature (Benevides & Lane, 2015; Zantinge et al., 2019). When considered together, these results suggest that increased heart rate and negative affect are observable markers of emotional dysregulation in infants later diagnosed with ASD and highlight the importance of selecting appropriate measures when examining complex psychological constructs.

Our study had several limitations. First, our results were derived from an IL sibling sample and thus may not be generalizable to children with non-familial ASD. Second, our clinical classifications at 24 months were completed virtually for a subset of our LL and IL participants because of COVID restrictions. Although clinical impression was based off of assessment and parent report, it is unclear how classification may have been impacted. Third, we did not collect height and weight measurements to control for BMI on the change in baseline heart rate from 6 to 18 months of age. Fourth, the diagnosis of ASD in the probands was not confirmed using a gold standard research assessment. Fifth, our sample of IL-ASD infants was relatively small, which may have underpowered our analyses. As such, these results should be corroborated in a larger sample of IL-ASD participants with a confirmed clinical diagnosis of ASD using gold standard methodology. Nevertheless, our study also has several strengths. We measured behavioral and physiological responses to positive and negative EE tasks at 6, 12, and 18 months and included three baseline periods to minimize affective and physiological carry-over.

Figure 4. Affect (A), gaze (B), and heart rate (C) responses during the Negative Tasks phase for the three groups by age.
effects between positive and negative tasks. Future research will consist of comparing affect, gaze, and physiological responses with a larger sample of children who IL-ASD defined by 36-month clinical best estimate diagnoses to validate and further elucidate differences between IL participants who receive a diagnosis of ASD and those who do not. The present study contributes to the growing literature indicating that ER may serve as an early biomarker of ASD vulnerability and may inform treatment strategies that could disrupt the developmental pathways between ER abnormalities and social, behavioral, and academic difficulties (Blair & RAZZA, 2007; EISENBERG et al., 2000, 2010; NOLAN et al., 2001; UPSHUR et al., 2009; WELSH et al., 2010).

Conclusion
Our findings indicate that affect, gaze, and physiological reactivity can differentiate between groups of children who are classified with ASD and their peers. Although these findings may not generalize to families who do not have an older child with ASD, the ability of our EE Task to identify differences early in development is similar to research using other technology-based approaches including electroencephalogram (EEG; EL-SABBAGH et al., 2012), magnetic resonance imaging (MRI; ELISON et al., 2013), and visual orienting (SACREY et al., 2014). As such, behavioral and physiological reactivity during emotionally evocative tasks may inform early intervention approaches for both IL and LL infants who show early signs of ASD.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0954579422001286

Acknowledgements. The authors thank the Canadian Institutes of Health Research, Brain Canada, Azrieli Foundation, and Autism Science Foundation for funding our research and Dr. Azadeh Kushki for her physiological expertise.

Funding statement. This work was supported by the Canadian Institutes of Health Research, Brain Canada, Azrieli Foundation, and Autism Science Foundation.

Conflicts of interest. None.

References