SPECIAL ISSUE

Developmental Psychobiology **WILEY**

Time–frequency dynamics of error monitoring in childhood: An EEG study

Santiago Morales[1,2] | Maureen E. Bowers[1] | Stephanie C. Leach[1] |
George A. Buzzell[3] | William Fifer[4] | Amy J. Elliott[5,6] | Nathan A. Fox[1]

[1] Department of Human Development and Quantitative Methodology, University of Maryland, College Park, Maryland, USA

[2] Department of Psychology, University of Southern California, Los Angeles, California, USA

[3] Department of Psychology, Florida International University, Miami, Florida, USA

[4] Department of Psychiatry, Columbia University, New York, New York, USA

[5] Avera Research Institute, Sioux Falls, South Dakota, USA

[6] Department of Pediatrics, University of South Dakota School of Medicine, Sioux Falls, South Dakota, USA

Correspondence
Santiago Morales, 501 Seeley G. Mudd Building University of Southern California, Los Angeles, CA 90089, USA.
Email: santiago.morales@usc.edu

Funding information
NIH Office of the Director, Grant/Award Number: UH3 OD023279

Abstract

Error monitoring allows individuals to monitor and adapt their behavior by detecting errors. Error monitoring is thought to develop throughout childhood and adolescence. However, most of this evidence comes from studies in late childhood and adolescence utilizing event-related potentials (ERPs). The current study utilizes time–frequency (TF) and connectivity analyses to provide a comprehensive examination of age-related changes in error-monitoring processes across early childhood ($N = 326$; 50.9% females; 4–9 years). ERP analyses indicated the presence of the error-related negativity (ERN) and error positivity (Pe) across all ages. Results showed no error-specific age-related changes in the ERN and the Pe. However, TF analyses suggested error-related frontocentral responses in delta and theta signal strength (power), delta consistency (intertrial phase synchrony), and delta synchrony (interchannel phase synchrony) between frontrocentral and frontolateral clusters—all of which increased with age. Additionally, the current study examines the reliability and effect size estimates of the ERP and TF measures. For most measures, more trials were needed to achieve acceptable reliability than what is commonly used in the psychophysiological literature. Resources to facilitate the measurement and reporting of reliability are provided. Overall, findings highlight the utility of TF analyses and provide useful information for future studies examining the development of error monitoring.

KEYWORDS
EEG, error monitoring, error positivity, error-related negativity, time frequency

1 | INTRODUCTION

Error monitoring, a component of cognitive control, allows individuals to monitor and adapt their behavior by detecting errors. In addition to being crucial for goal-directed behavior, error monitoring has been linked to a variety of developmental processes and outcomes across development, including executive function (Grammer et al., 2018), academic performance (M. H. Kim et al., 2016; S. H. Kim et al., 2020), and externalizing and internalizing psychopathology (Fox et al., 2020; Hajcak et al., 2003; Pasion & Barbosa, 2019; Shackman et al., 2011). Like other aspects of cognitive control, error monitoring is thought to

develop throughout childhood and adolescence (Buzzell et al., 2017; Davies et al., 2004; Tamnes et al., 2013). However, most of this evidence comes from studies in late childhood and adolescence utilizing event-related potentials (ERPs).

In the current study, we utilize a Go/No-Go task specifically designed for young children along with time–frequency (TF) and connectivity analyses of electroencephalography (EEG) data to provide a comprehensive examination of age-related changes in error-monitoring processes across childhood. Moreover, the reliability of these error-related neurocognitive measurements are unknown in young children, potentially limiting their application as markers of

cognitive control and early risk factors for other developmental outcomes (e.g., psychopathology). As such, the current study also provides internal consistency reliability estimates and effect sizes of these error-related measures at different ages in a large sample of children across childhood (4–9 years).

1.1 | ERPs: error-related negativity and error positivity

The error-related negativity (ERN) is a frontocentral negative deflection approximately 0–100 ms after an incorrect response on a task (Falkenstein et al., 1991; Gehring et al., 1993). EEG source modeling and functional magnetic resonance imaging (fMRI) studies suggest that the ERN is generated, at least in part, within the anterior cingulate cortex (ACC)—a brain region that integrates threat, pain, and negative feedback to guide future behavior (Buzzell et al., 2017; Shackman et al., 2011). The error positivity (Pe) is a centroparietal positive deflection approximately 200–500 ms after an error (Falkenstein et al., 2000). Although the functional significance of the Pe is still unclear, the Pe is thought to reflect the awareness of having made an error (Boldt & Yeung, 2015; Nieuwenhuis et al., 2001; Overbeek et al., 2005; Steinhauser & Yeung, 2010).

In line with other developmental increases in cognitive control (Luna et al., 2004; Zelazo et al., 2013), studies examining the development of error-related ERPs suggest that the ERN and Pe increase in magnitude with age (Buzzell et al., 2017; Davies et al., 2004; Ladouceur et al., 2007; Tamnes et al., 2013). However, most of this evidence comes from studies in late childhood and adolescence. Emerging studies with younger children suggest a similar trend (DuPuis et al., 2015; Torpey et al., 2012), but evidence is more equivocal (Grammer et al., 2014; Lo et al., 2015). Therefore, more data with large samples of young children are needed to better understand the development of error monitoring across childhood.

1.2 | TF analyses

Importantly, most studies examining the development of error monitoring using EEG have limited their analyses to ERPs. ERPs do not fully leverage the information contained in the EEG signal, which may help explain the inconsistent findings. ERP analyses assume that the component of interest is temporally synchronous across trials, only focusing on neural activity that is time-locked to the event of interest and disregarding non-phase-locked signals (Luck, 2014). In contrast, TF analyses measure the amplitude and phase of neural oscillations across different frequencies. By differentiating between amplitude and phase information, both phase-locked and non-phase-locked signals can be studied in relation to an event of interest. Moreover, TF measures provide more direct information regarding the neurophysiological mechanisms underlying error monitoring that is present in the EEG data. For example, TF analyses can provide distinct measures of signal strength (power), consistency of phase oscillations across trials (intertrial phase

synchrony; ITPS), and consistency of phase oscillations across different electrodes across trials (interchannel phase synchrony; ICPS) (Cohen, 2014). Consequently, TF analyses of error-related processes may provide unique insight into the neurocognitive processes underlying error monitoring and the specific neural mechanisms that change across development.

TF analyses of error-related processes have primarily focused on the theta frequency band (Buzzell et al., 2019; Cavanagh & Shackman, 2015; Narayanan et al., 2013; Steele et al., 2016). In support for TF analyses providing a unique perspective on the development of error monitoring, the limited TF data suggest distinct developmental changes in power and phase synchrony that cannot be observed by only examining the ERP components (Bowers et al., 2018; DuPuis et al., 2015). For instance, in a longitudinal sample from 5 to 7 years, DuPuis and colleagues (2015) found different developmental trajectories of power and phase synchrony—such that power decreased with age, whereas phase synchrony increased with age. Moreover, phase synchrony mediated the relationship between age and ERN amplitude (DuPuis et al., 2015). Studies examining the variability of the ERN also suggest increases in the consistency of the signal with age (Gavin et al., 2019).

Error-related functional connectivity measures (e.g., ICPS) have not been reported in young children, but work in adults suggests increases in theta ICPS connectivity between frontocentral and frontolateral sites after errors (Cavanagh & Frank, 2014). This pattern of connectivity is thought to reflect frontocentral structures (e.g., ACC) recruiting lateral regions such as dorsolateral prefrontal cortex (dlPFC) or motor regions, involved in posterror changes in attention and behavior (Buzzell et al., 2019). In a sample of adolescents, Buzzell and colleagues (2019) found similar patterns of error-related theta-band connectivity (ICPS) to those observed with adults. However, studies examining developmental changes in error-related connectivity across childhood are nonexistent.

1.3 | Current study

In the current study, we utilize ERP and TF approaches in a Go/No-Go paradigm adapted for young children to provide a comprehensive analysis of error-related processes across childhood (4–9 years). For ERPs, we expected to observe the ERN and the Pe in response to errors. Based on previous studies, we hypothesized age-related increases in ERN and Pe magnitudes; namely, more negative ERN and more positive Pe.

Although most studies examining error-related processes have focused on theta band frequencies, neuronal oscillations in the delta range have also been implicated in error-related processes (Cavanagh et al., 2009; Munneke et al., 2015). Moreover, delta power and ITPS significantly increase with age (Bowers et al., 2018). Thus, we examined power, ITPS, and ICPS in both the delta and theta band. We expected to observe increased frontocentral theta power and ITPS on error, relative to correct trials. We hypothesized age-related increases in these error-related effects in delta and theta power and ITPS. For

connectivity, we expected to observe increased delta and theta band ICPS between frontocentral and frontolateral regions for error, relative to correct trials. Additionally, we hypothesized that this connectivity would increase with age.

The overall goal of the current study is to provide a comprehensive description of error-related measures across childhood to inform the development of cognitive control and provide metrics that can predict other outcomes (e.g., neurodevelopment or psychopathology). The validity of these error-related measures as individual-level predictors of other outcomes depend on their reliability. However, the reliability of these error-related neurocognitive measurements are unknown in young children. Therefore, in addition to examining age-related changes in ERPs and TF measures, we also provide estimates of reliability and effect sizes at varying numbers of trials, separately at different ages.

2 | METHODS

2.1 | Participants

Children ($N = 566$; $M_{age} = 5.97$ years; 297 females) were recruited as part of a large study examining the impact of early environmental exposures on children's development between September 2018 and March 2020. Participants were all part of the Prenatal Alcohol in SIDS and Stillbirth Safe Passage (PASS) cohort. More details on the larger study, including recruitment has been previously published (Dukes et al., 2014). The data utilized in the current study were collected as part of the Environment Influences on Child Health Outcomes (ECHO) study PASS Cohort in South Dakota (Gillman & Blaisdell, 2018). The EEG assessments in the current study involve a cross-sectional sample of children at 4, 5, 7, and 9 years of age.

Of the 550 participants that agreed to complete the Go/No-Go task with EEG: two completed the task but were removed due to technical difficulties, 55 were removed because they did not complete at least 160 trials or had lower than 60% overall accuracy during the task, and 167 did not have sufficient trials per condition (at least six artifact free-trials), leaving a total of 326 (50.9% females) participants included in the present study. The sample who provided usable EEG data was largely White (80.7%), followed by American Indian (12.0%), and other (7.4%). Mothers reported on average 15.0 years of education (Range = 7–17 years) when they joined the study. Additionally, participants had the following years of age distribution: 4 ($n = 58$; $M_{age} = 4.24$; $SD_{age} = 0.15$), 5 ($n = 97$; $M_{age} = 5.23$; $SD_{age} = 0.16$), 7 ($n = 119$; $M_{age} = 7.20$; $SD_{age} = 0.13$), and 9 ($n = 52$; $M_{age} = 9.25$; $SD_{age} = 0.13$). Children that contributed EEG data did not differ in sex ($p = .35$) or birthweight ($p = .67$), but were more likely to be older ($p < .001$), such that 45% of the 4-year olds, 61% of the 5-year olds, 64% of the 7-year olds, and 69% of the 9-year olds contributed data. Moreover, children who contributed data were more likely to be White ($p = .005$) and have more educated mothers ($p < .001$) than children who did not contribute usable EEG data.

2.2 | Protocol/procedure

Prior to data collection, all primary caregivers of the participants provided informed consent and all children provided assent. As part of a larger laboratory visit including several health (e.g., anthropometrics and spirometry) and cognitive assessments, participants completed a Go/No-Go task called the "Zoo Game" while EEG was measured. Before the Zoo Game, participants completed a 3-min baseline and a three-stimulus auditory oddball (not reported here). Participants were seated approximately 70 cm in front of the presentation computer. The task was presented in E-Prime 2.0.10 (Psychology Software Tools, Pittsburgh, PA) and is publicly available here: https://github.com/ChildDevLab/Tasks. At the end of the visit, families were compensated and children were provided a small toy. The Avera Institutional Review Board approved all study procedures.

2.3 | EEG data acquisition

EEG was recorded using a 64-channel HydroCel Geodesic Sensor Net and sampled at 500 Hz via EGI software (Net Station Version 5.4; Electrical Geodesics, Inc., Eugene, OR). The nets had the four face channels (E61–E64) removed to measure other psychophysiological measures (e.g., HR) and were not used to collect EEG. Prior to data collection, impedance values were checked for all EEG channels and confirmed to be below 50 kΩ.

2.4 | Zoo Game (Go/No-Go task)

The Zoo Game is a computer-based Go/No-Go task developed by Fox and colleagues (He et al., 2010; Morales et al., 2020; Troller-Renfree et al., 2019) based on the design of Durston et al. (2002). The task was further adapted by Grammer et al. (2014) to be used for collecting ERPs by increasing the number of trials and changing the task parameters (e.g., intertrial interval). Children were instructed to help a zookeeper catch animals in a zoo that had escaped from their cages. Additionally, children were told that the orangutan was the zookeeper's assistant helping to catch the animals, so they should not catch the orangutan. At the age of 4 years, there was only one picture of an orangutan. At the other ages, there were three different pictures of orangutans. Children were instructed to press a button to catch all of the animals (Go trials) but to withhold responses for the orangutan (No-Go trials). On each trial, an animal stimulus was presented on the screen for 750 ms, followed by a blank screen for 500 ms or until the child responded; the intertrial interval was jittered between 500 and 1000 ms—which differs from Grammer et al. (2014) who used 200–300 ms. Children practiced the task until the experimenter believed they understood the task, then completed up to 320 test trials, presented in eight blocks of 40 trials each. The task consisted of 75% Go trials and 25% No-Go trials. As in previous studies involving RT-based tasks with children (Bowers et al., 2021; Morales

et al., 2016), all Go and No-Go data were cleaned to remove anticipatory responses (RTs < 150 ms) prior to the computation of accuracy measures. Response accuracy was calculated on both Go and No-Go trials.

2.5 | EEG processing

EEG data were processed using a standardized preprocessing and postprocessing pipeline specifically designed for developmental data (MADE; Debnath et al., 2020; Leach et al., 2020). In brief, EEG data were exported to Matlab (Mathworks, Natick, MA) for offline processing with EEGLAB (v13.6.5b) toolbox (Delorme & Makeig, 2004) and in-house Matlab scripts (Matlab 2014). Continuous data were high pass filtered at 0.3 Hz and then low pass filtered at 49 Hz. Artifact-laden channels were identified and removed using the EEGLAB plug-in FASTER (Nolan et al., 2010). To further remove ocular and muscle artifacts, extended independent component analysis (ICA) was performed on an identical copy of the dataset. In order to improve ICA decomposition, this copied dataset was high pass filtered at 1 Hz and segmented into 1 s epochs. Then, noisy segments of the data and EMG-like activity were rejected using a voltage threshold of ± 1000 μV and spectral threshold (range -100 dB to $+30$ dB) within the 20−40 Hz frequency band. If a channel had an identified artifact in more than 20% of the epochs, that channel was removed from both the ICA copied dataset and the original dataset. ICA was then run on the copied dataset, and ICA weights were subsequently applied back to the original dataset (Debner et al., 2010). Artifactual components were then removed from the original dataset by an automated process that included using the Adjusted-ADJUST algorithm (Leach et al., 2020). Data were then epoched into 3000 ms segments that started 1000 ms before the response onset.

After ICA artifact removal and epoching, a two-step procedure for identifying residual artifacts was employed. First, any epochs where ocular channel (E1, E5, E10, and E17) voltages exceeded ± 150 μV were rejected to remove residual ocular activity not removed through ICA. Second, for any epoch in which nonocular channel voltages exceeded ± 125 μV, these channels were interpolated at the epoch level. If greater than 10% of the channels (not considering globally rejected channels) exceeded ± 125 μV in the epoch, the epoch was rejected. All missing channels were interpolated using a spherical spline interpolation and then the data were referenced to the average of all the electrodes. The average number of interpolated channels per epoch (including those globally rejected) was 4.80% (range 1.3–13.2%). Finally, all trials with anticipatory responses (RTs < 150 ms) were removed.

The average number of remaining trials included for analysis for each of the two conditions was 161.46 for correct and 16.31 for error trials. Based on cutoffs determined by previous studies (Pontifex et al., 2010; Steele et al., 2016), children with less than six trials for any condition were removed from further analysis (see above).

2.5.1 | ERPs

ERPs for each child were averaged separately for each condition and baseline corrected to the average voltage in the -300 to -100 ms prestimulus period, using mean amplitude measures for each child and condition. The time windows and electrode clusters for each ERP component (ERN and Pe) were selected based on previous publications with developmental populations (Brooker, 2018; Grammer et al., 2018) and visual inspection. For the ERN, we focused on the frontrocentral electrode cluster (FCz; E4, E7, and E54) 0–100 ms after the response. For the Pe, we focused on the parietal electrode cluster (E33, E34, E36, and E38) 200–500 ms after the response.

2.5.2 | TF power

TF power in each epoch of interest was computed using custom MATLAB scripts (Morales & Bowers, 2021), adapted from Cohen (2014). First, a surface Laplacian filter was applied to the epoched data to mitigate volume conduction over the scalp by filtering out spatially broad features of the data (Cohen, 2014) in order to improve both spatial and functional specificity of brain activity (Tenke & Kayser, 2012). Each CSD converted epoch was convolved with Morlet wavelets, which estimated spectral power in the frequency range 1−30 Hz (in 60 steps spaced logarithmically). To optimize the TF resolution, wavelet cycles were set at three cycles at the lowest frequency (1 Hz) increasing to 10 cycles at the highest frequency (30 Hz). Power was computed separately for both conditions (error and correct) for all channels. For each condition, power was normalized using a (dB) transform (dB power = $10 \times$ log 10[power/baseline]), where the baseline was the average power for that condition from -300 to -100 ms window before the response onset (Delorme & Makeig, 2004). Finally, data were downsampled to 50 Hz to facilitate data manipulation.

2.5.3 | ITPS and ICPS

ITPS and ICPS were computed using custom MATLAB scripts (Morales & Bowers, 2021) based on Cohen (2014). ITPS measures the consistency of phase oscillations for each frequency window, at each timepoint across trials. ICPS measures the consistency of phase oscillations between two channels (or clusters of channels) for each frequency window, at each timepoint across trials. ITPS and ICPS values range from 0, indicating random phase alignment at that time point, to 1, representing perfect phase alignment at that time point. ITPS was calculated by taking the phase angle difference across trials, then averaging the phase angle differences. ICPS was calculated by taking the phase angle difference between FCz (E4) and electrodes from two frontolateral channel clusters (left E12 and E13; right E59, and E60), averaging the phase angle difference between each electrode pair across trials, then averaging the phase angle differences between each electrode pair. A subsampling procedure was used when estimating

ICPS and ITPS to eliminate biases associated with having different numbers of trials per condition (Cohen, 2014). ITPS and ICPS was calculated for five trials that were randomly selected per condition. This subsampling procedure was performed 100 times to ensure all the data were used, then all subsamples were averaged. ITPS and ICPS were baseline corrected per condition based on −300 to −100 ms before stimulus onset. This process created ITPS and ICPS surfaces per condition with the same dimensions as the TF measures for each electrode for each participant. Finally, data were downsampled to 50 Hz to facilitate data manipulation.

2.5.4 | TF regions of interest selection

The frequencies and time windows for the regions of interest (ROIs) were selected a priori based on previous literature on developmental populations (Buzzell et al., 2019, 2020) and confirmed with visual inspection. Although the alpha peak increases from infancy to childhood, by 4 years of age most children display their alpha peak at 9 Hz (Marshall et al., 2002; Thorpe et al., 2016). Because of this and in line with previous studies with young children (3–6 years; e.g., Canen & Brooker, 2017; DuPuis et al., 2015; Kikuchi et al., 2011; Orekhova et al., 2006), we chose the theta frequency band as 4–8 Hz. Delta was defined as the frequencies below theta (1–4 Hz). We focused on delta and theta power and ITPS over the frontocentral electrodes (FCz; E4, E7, and E54) and ICPS between FCz and frontolateral channel clusters (left E12 and E13; right E59, and E60). To determine the time windows of interest unbiased for condition effects, theta and delta power, ITPS, and ICPS were each averaged over all conditions throughout the interval from −200 to 600 ms post stimuli. As anticipated, we observed a clear delta and theta response between 0 and 300 ms for power, ITPS, and ICPS. Examining topographic maps and TF surfaces (see online supplement) supported the selections of these ROIs.

2.6 | Analytic approach

In order to examine condition effects and if these effects significantly changed with age, we performed a set of multilevel models (MLMs) using the *nlme* package in R (Pinheiro et al., 2016; R Development Core Team, 2008). Models were performed separately for each outcome of interest; namely, each ERP (ERN and Pe), power, ITPS, ICPS for delta and theta frequencies. Models were examined in two steps. In the first step, we included the effect of condition (error vs. correct) with correct as the reference condition in order to test the main effect of condition. In the second step, we included the main effect of age and the interaction between condition and age (condition × age). A significant interaction would indicate that the effect of errors changes with age. This would be equivalent to the difference score between error and correct being significantly associated with age. The motivation for this approach (over just showing the difference score) is that it

tests age-related changes in the error-related trials (e.g., ERN) and the difference between error and correct trials (e.g., delta ERN), as both measures are commonly used in the literature. Moreover, all models controlled for the effects of gender, number of trials in each condition, and data collection site and included the random intercept effect for each participant.

2.7 | Reliability and effect sizes

2.7.1 | ERP

In order to estimate the internal consistency reliability and effect size of each measure, we followed an approach similar to one previously described (Leach et al., 2020; Towers & Allen, 2009). In short, this approach utilizes the trial-by-trial ERP estimates to identify the minimum number of trials needed to obtain a reliable measure and observe a significant condition effect (e.g., ERN). For this, reliability and effect size estimates were obtained for an increasing number of trials from 4 to 32 trials in steps of four trials. Because estimates heavily depend on which trials are sampled, for each number of trials, *n*, 3000 iterations of split-half correlations were calculated. For each iteration, *n* trials were randomly selected from all available trials for that condition for a given participant, provided there were enough trials for the given *n*. For effect size estimates, Cohen's d was computed for each measure by contrasting error and correct trials. For reliability estimates, the selected trials were then halved by randomly assigning trials to one of the halves. The internal consistency reliability estimates were measured via the Spearman–Brown split-half correlation method.

This process generated 3000 effect size and reliability estimates for a given *n* for each measure. This provided us not only with a more robust estimate for a given trial number (*n*), but also an estimate of the variability around that estimate (i.e., a measure of the resampling distribution). Thus, we present the average and 95% confidence intervals (CIs) from the resampling distribution for each measure separately for each age. Finally, in order to provide a measure of overall reliability (i.e., including all available participants and trials), we estimated the reliability using the Spearman–Brown split-half correlation method described above (3000 subsamples), but using all of the available trials.

2.7.2 | TF power

Because the baseline normalization of TF power is nonlinear, we did not use trial-by-trial estimates. We followed the split-half approach with subsampling described above for ERPs, but the TF power estimates for a given half with number of trials, *n*, were normalized using a dB transform. The baseline activity was taken as the average power of those trials at each frequency band for each condition from −300 to −100 ms prestimulus. The rest of the procedure to calculate reliability and effect

size estimates for increasing numbers of trials and overall reliability were the same as the one described above for ERPs.

2.7.3 | ITPS and ICPS

Because ITPS and ICPS estimates are defined across trials, it is not possible to follow the trial-by-trial approach described above. Reliability and effect size estimates for ITPS and ICPS were also calculated in increasing numbers of trials (4–32 in steps of four trials). For each number of trials, n, the randomly sampled trials were split in half and ITPS and ICPS were calculated separately for each half. To ease computation time, data were downsampled to 50 Hz before estimating ITPS and ICPS. The internal consistency reliability estimates were measured via the Spearman–Brown split-half correlation method. This was repeated 100 times and the ITPS and ICPS reliability estimates were averaged to try to mitigate the variability inherent in subsampling random trials and randomly splitting those trials into two halves. For plotting, negative reliability estimates were set to zero.

Notably, when iterating across number of trials (from 4 to 32), the number of participants included in the effect size and reliability estimates decreased as the number of trials increased. Only reliability and effect size estimates with a minimum of six participants are presented (Leach et al., 2020).

3 | RESULTS

3.1 | ERPs

3.1.1 | ERN

As hypothesized and shown in Figure 1, we observed a clear ERN component, indexed by a more negative amplitude deflection between 0 and 100 ms in frontocentral electrodes in response to errors, compared to correct responses, $b = -5.97$, $p < .001$. Examining the interaction between condition and age revealed that the response to errors, compared to correct responses, did not change with age, $b = 0.03$, $p = .849$, but there was a significant main effect of age, $b = -0.72$, $p < .001$.

3.1.2 | Pe

As expected and shown in Figure 1, we observed a Pe component, indexed by an amplitude increase between 200 and 500 ms in parietocentral electrodes in response to errors, compared to correct responses, $b = 13.41$, $p < .001$. The interaction between condition and age revealed that the response to errors, compared to correct responses, did not significantly change with age, $b = 0.58$, $p = .057$, but there was a significant main effect of age, $b = 1.07$, $p = .004$ (Figure 1).
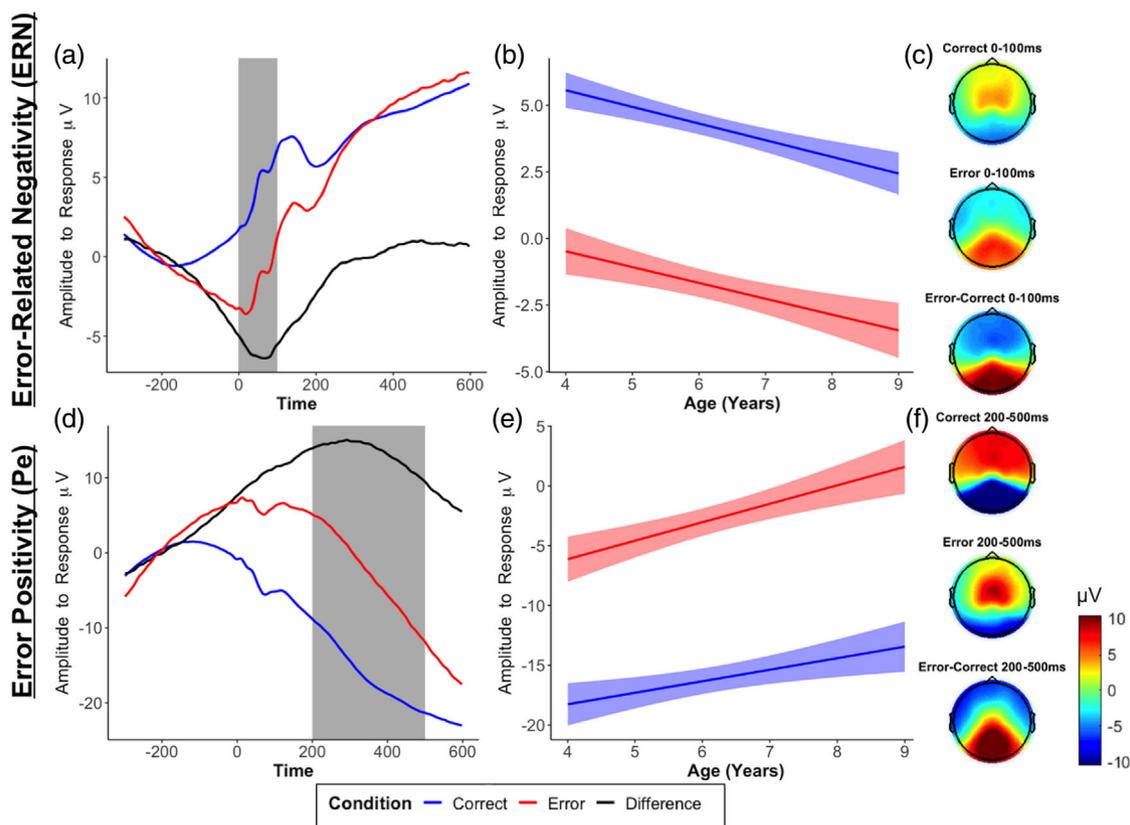
FIGURE 1 Error-related ERP measures: ERN and Pe. Average waveforms across all participants for the ERN (FCz; a) and the Pe (Pz; d), age-related changes in ERPs in the selected time window (b and e), and topographic plots (c and f) for each condition and condition differences
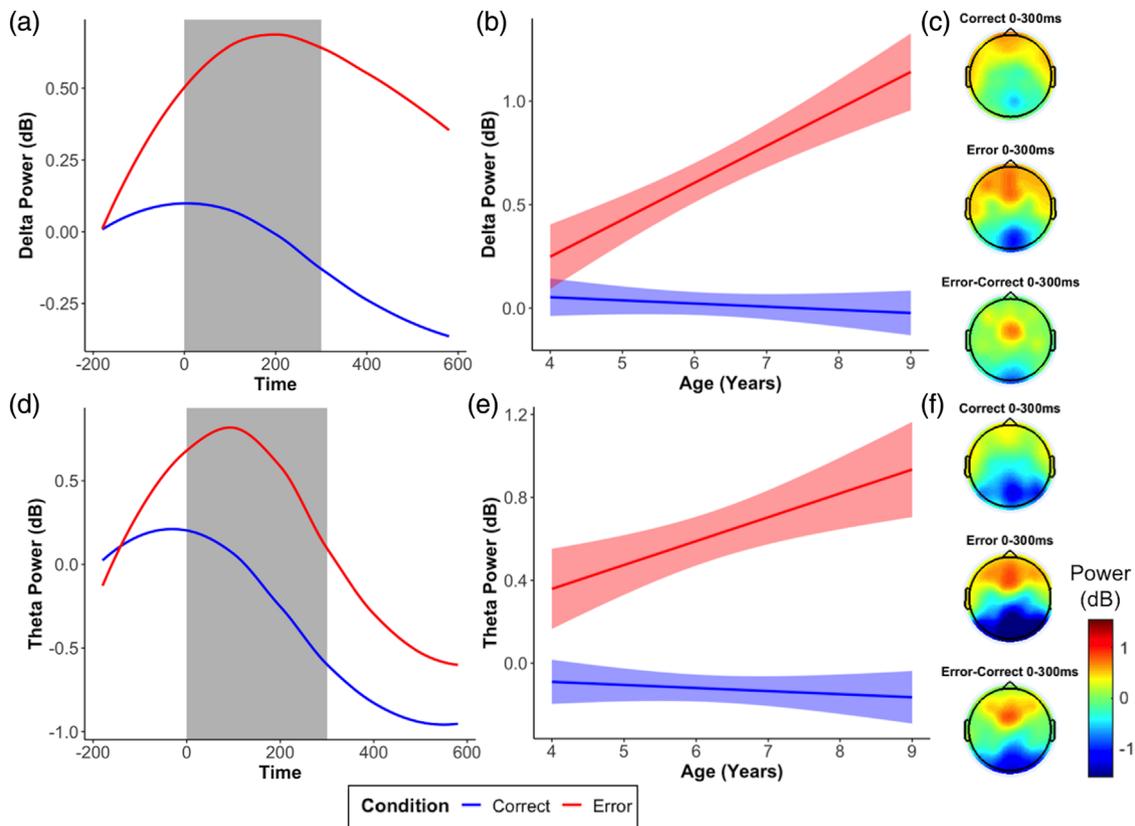
FIGURE 2 Time–frequency dynamics of delta and theta power during error and correct trials at the FCz cluster. Plots show time–frequency power for each condition across all participants (a and d), age-related changes in time–frequency power in the selected time window (b and e), and topographic plots (c and f) for each condition

3.2 | TF power

As hypothesized and shown in Figure 2, we observed delta and theta power increases between 0 and 300 ms in frontocentral electrodes in response to errors, compared to correct responses, $b = 0.62$, $p < .001$ for delta and $b = 0.73$, $p < .001$ for theta. There were also significant interactions between condition and age for delta, $b = 0.19$, $p < .001$, and theta, $b = 0.13$, $p < .001$. As shown in Figure 2, both the delta and theta responses to errors, compared to correct responses, significantly increased as age increased.

3.3 | ITPS

We also observed delta ITPS increases between 0 and 300 ms in frontocentral electrodes in response to errors, compared to correct responses ($b = 0.03$, $p < .001$; Figure 3). Adding the interaction with age revealed a significant interaction for delta ITPS ($b = 0.007$, $p < .001$), suggesting that delta ITPS to errors, compared to correct responses, significantly increased as age increased (Figure 3). However, for theta ITPS, we did not observe either a significant effect of condition ($b = -0.005$, $p = .122$), age ($b = 0.001$, $p = .325$), or an interaction with age ($b = -0.004$, $p = .057$).

3.4 | ICPS

As predicted and shown in Figure 4, we observed delta ICPS increases between 0 and 300 ms between a frontocentral electrode and frontolateral electrodes in response to errors, compared to correct responses, $b = 0.016$, $p < .001$. The interaction with age revealed a significant interaction predicting delta ICPS, $b = 0.004$, $p = .015$, showing that delta ICPS to errors, compared to correct responses, significantly increased as age increased (Figure 4). For theta ICPS, we observed a significant effect of condition ($b = 0.006$, $p = .032$), but we did not observe a significant effect of age ($b = -0.001$, $p = .326$) or an interaction with age ($b = 0.002$, $p = .155$).

3.5 | Reliability and effect sizes

3.5.1 | ERPs

Figure 5 shows the reliability estimates by increasing number of trials for the two ERPs (ERN and Pe) in each condition. In Table 1, we show the minimum number of trials needed on average and within the 95% CI to achieve acceptable reliability for each age. In general, older ages seemed to have better reliability and require fewer trials to achieve
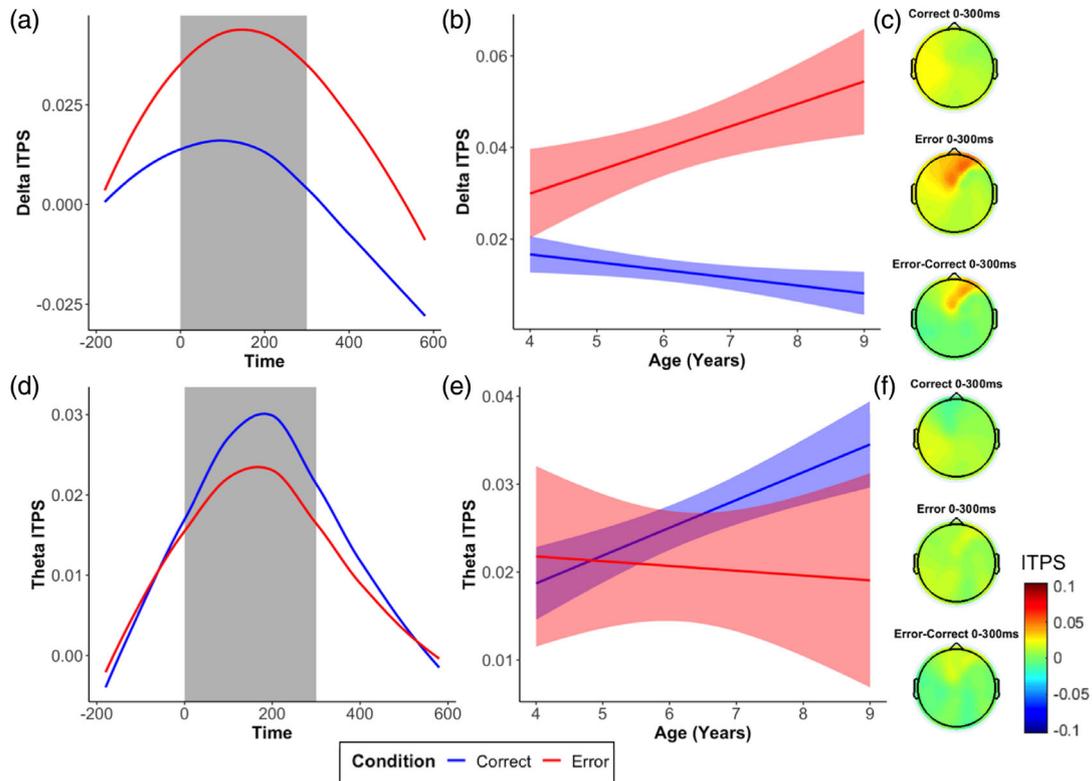
FIGURE 3 Time–frequency dynamics of delta and theta intertial phase synchrony (ITPS) during error and correct trials at the FCz cluster. Plots show ITPS for each condition across all participants (a and d), age-related changes in ITPS in the selected time window (b and e), and topographic plots (c and f) for each condition
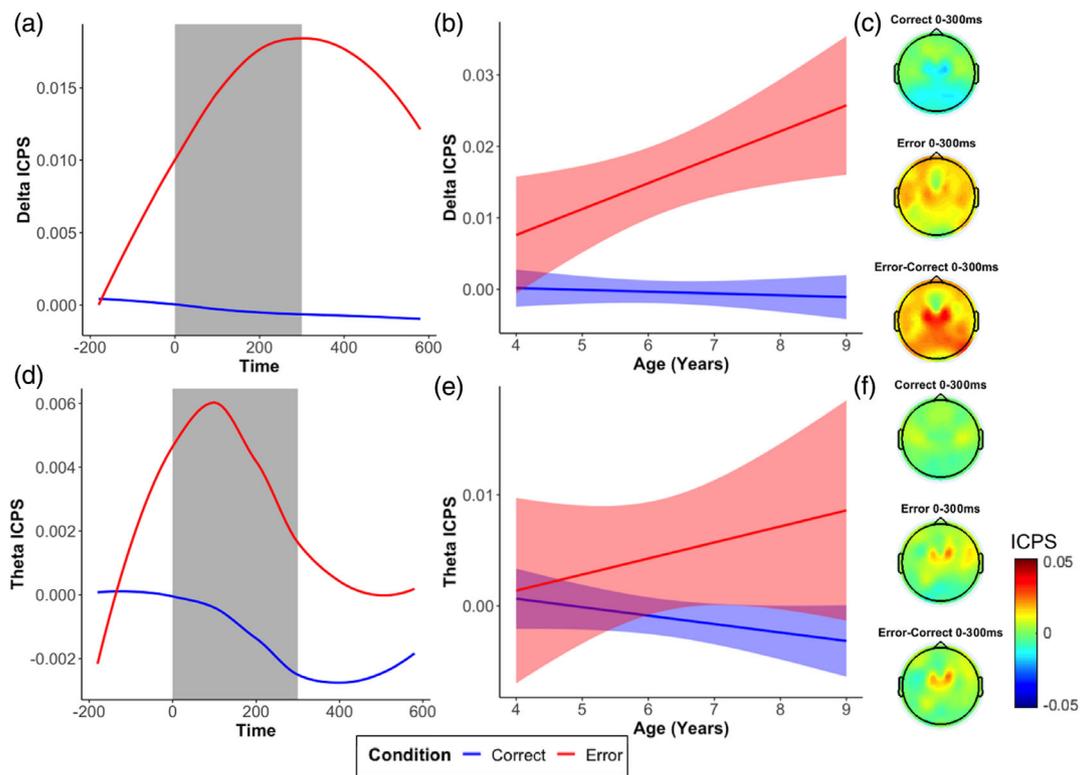
FIGURE 4 Time–frequency dynamics of delta and theta interchannel phase synchrony (ICPS) during error and correct trials at frontolateral clusters. Plots show ICPS for each condition across all participants (a and d), age-related changes in ICPS in the selected time window (b and e), and topographic plots for each condition (c and f)
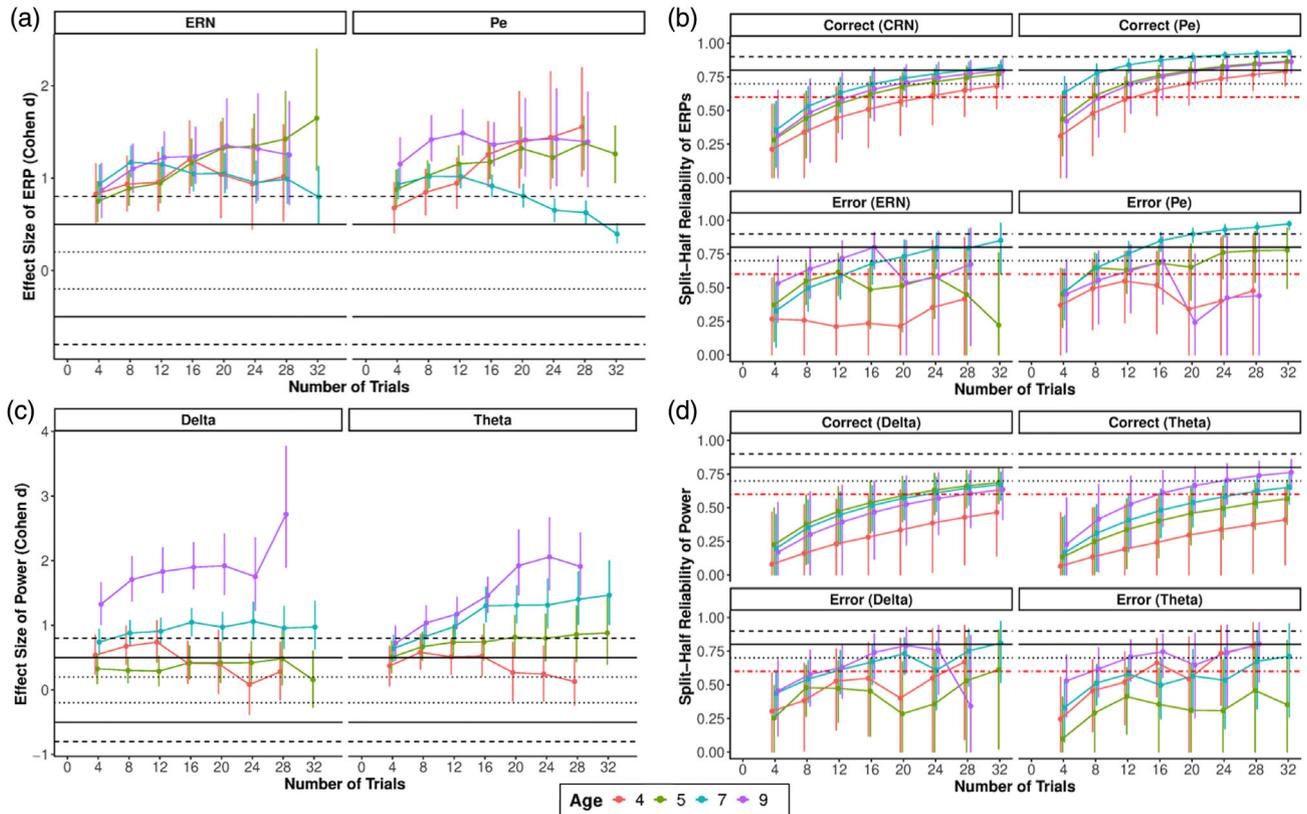
FIGURE 5 ERP and time–frequency power effect size and reliability estimates across increasing numbers of trials for each age. Plots show the effect size of error versus correct for each ERP (a) and power at each frequency (c), as well as split-half reliability for each ERP (b) and power at each frequency (d) for a given condition

acceptable reliability. Moreover, the Pe seemed to have better reliability than the ERN. This was confirmed when examining the overall reliability (Table 1). For effect sizes, both the ERN and the Pe showed large effect sizes (>.8), which did not show a clear pattern across trials or age.

3.5.2 | TF power

Figure 5 shows the reliability estimates by increasing number of trials for the TF power estimates and their corresponding baseline periods. In Table 1, we show the minimum number of trials needed on average and within the 95% CI to achieve acceptable reliability for each age. In general, older ages seemed to have better reliability and require fewer trials to achieve acceptable reliability. This was confirmed when examining the overall reliability (Table 1). For effect sizes, as trials increased, older ages reached higher effect sizes when comparing error to correct responses in delta and theta power. On the other hand, error versus correct effect sizes remained relatively low at younger ages around a small effect size, regardless of how many trials were included.

3.5.3 | ITPS

Figure 6 shows the reliability estimates by increasing number of trials for the ITPS estimates for each condition. In Table 1, we show the min-

imum number of trials needed on average to achieve acceptable reliability for each age. Note that because of the limited subsampling, the estimates across trials vary more between numbers of trials. In general, it showed the expected pattern of increasing reliability with increasing numbers of trials and older ages seemed to have better reliability and require fewer trials to achieve acceptable reliability. For estimates of effect size, there is no clear pattern with increasing trials. For delta ITPS, several age groups display a medium to large effect size. However, this was not the case for theta ITPS, which showed no clear age effect. Finally, because of the lack of trial-level data, there are no CIs and we do not draw any conclusions between ages and conditions.

3.5.4 | ICPS

Figure 6 shows the reliability estimates by increasing number of trials for the ICPS estimates for all conditions. In Table 1, we show the minimum number of trials needed on average to achieve acceptable reliability for each age. Note that because of the limited subsampling, the estimates across trials vary more between numbers of trials. Descriptively, it showed the expected pattern of increasing reliability with increasing numbers of trials, but for correct trials for all ages and younger ages for errors, acceptable reliability was not reached within 32 trials. For estimates of effect size, there is no clear pattern with increasing trials. For delta ICPS, older age groups displayed larger effect sizes. However,

this was not the case for theta ICPS, which showed no clear age-related effect. Finally, because of the lack of trial-level data, there are no CIs and we do not draw any conclusions between ages and conditions.

4 | DISCUSSION

Error monitoring is an important component of cognitive control that has been linked to a variety of developmental processes and outcomes, including executive function (Grammer et al., 2018), academic outcomes (M. H. Kim et al., 2016; S. H. Kim et al., 2020), and psychopathology (Fox et al., 2020; Hajcak et al., 2003; Pasion & Barbosa, 2019; Shackman et al., 2011). Thus, it is important to better understand the development of error monitoring as well as the psychometric properties of the indices commonly used to measure it. In the current study, we utilized ERP and TF approaches in a Go/No-Go paradigm adapted for young children to provide a comprehensive analysis of error-related processes across childhood (4–9 years). In addition, the current study examined the reliability and effect size estimates of these ERP and TF measures.

4.1 | ERPs: ERN and Pe

Our ERP results showed that, as hypothesized, we observed error-related effects in the time windows and topography corresponding to the ERN and Pe components. This suggests that error-monitoring processes, including those associated with automatic error detection (ERN) and the awareness of having made an error (Pe; Nieuwenhuis et al., 2001), are present in early childhood. On the other hand, contrary to our expectations, the ERN (compared to the CRN) and the Pe (compared to the Pe for correct trials) did not significantly change with age. Importantly, we did observe the expected increases in the ERN and the Pe, but this effect was not exclusive to error trials and was also observed for correct trials (e.g., CRN). Several previous studies reporting significant age-related changes in ERN did not examine the specificity of these changes with respect to the CRN. Moreover, the lack of age-related changes may be due to the relatively limited age range in the current study compared to other studies or the fact that we are examining an earlier developmental period than most studies (Buzzell et al., 2017; Davies et al., 2004; Ladouceur et al., 2007; Tamnes et al., 2013). Future longitudinal studies in early childhood are needed to clarify the early development of error-related processes. However, it is also important that future studies examine the specificity of developmental changes to error-related processes rather than action monitoring in general.

4.2 | Time-frequency analyses: power, ITPS, and ICPS

In addition to our time-domain analyses, we examined several TF measures to differentiate between amplitude and phase information in the delta and theta frequencies. For TF power, we found error-related frontocentral responses in delta and theta. This frontocentral delta/theta increase in response to errors is thought to reflect the activation in the action monitoring system (putatively ACC) that functions as an "alarm" to detect errors and increases attention to instantiate control (Cavanagh & Frank, 2014). Moreover, we found that the strength of the signals after errors in both the theta and delta frequencies significantly increased with age suggesting important development in the error-monitoring system across childhood.

ITPS increased during error responses compared to correct responses in delta. Contrary to our initial hypotheses, this was not the case for ITPS in theta. In previous studies with adults, theta ITPS is generally more prominent after errors compared to correct responses (Aviyente et al., 2011; Cavanagh et al., 2009). In children and adolescents (7–18 years), there is evidence of a theta ITPS response after errors, compared to a preresponse baseline; however, it was not compared to correct responses (DuPuis et al., 2015; Gavin et al., 2019). Here, we show that young children do not show the expected error-related theta ITPS responses when compared to correct responses. This finding is similar to findings in a feedback task that found no difference in children (ages 10–12) when comparing loss to win, but condition differences seem to emerge at older ages (Crowley et al., 2014). However, in our age range, we did not observe age-related changes in theta ITPS, possibly due to a lack of differences between error and correct responses. In contrast, we did observe greater phase synchrony to error responses compared to correct responses in the delta frequency, another frequency whose dynamics are implicated in error processing. Indeed, visual inspection of data with adults also show significant ITPS in delta, perhaps even more prominently than theta, but this was not formally tested or reported in those studies (e.g., Cavanagh et al., 2009). Additionally, this delta ITPS response to errors, compared to correct responses, increased with age. This is in line with what was found with feedback-related delta ITPS (Bowers et al, 2018). These findings suggest that the development of error processing in childhood is marked by increased delta consistency in response to errors.

To our knowledge, this study is the first to investigate error-related frontocentral to frontolateral ICPS in younger children. In adults and adolescents, error-related frontocentral to frontolateral theta ICPS response in adults and adolescents is thought to reflect connectivity between the mediofrontal cortex (including the ACC) with dlPFC and motor areas, involved in instantiating control (Buzzell et al., 2019; Cavanagh et al., 2009). As expected, we observed increases in theta ICPS, but similar to the pattern of results for ITPS, we did not observe age-related changes in theta ICPS. In one of the few developmental studies of stimulus-related ICPS, no age-related changes in theta ICPS were reported across adolescence (Bowers et al., 2021). On the other hand, we found a strong delta ICPS response to errors when compared to correct responses in young children, suggesting that delta phase synchrony between frontocentral and frontolateral areas may play an important role in error processing in young children. Moreover, delta ICPS to errors exclusively increased with age, showing that delta-based synchrony

TABLE 1 Descriptives and minimum and overall reliability for ERP and TF power, ITPS, and ICPS measures

Age	N	Correct					Error				
		Mean	Min avg	Min CI	Total	% Data rejected	Mean	Min avg	Min CI	Total	% Data rejected
ERN											
4	58	5.31	24	8	.88	23.1	−1.14	–	16	.55	31.6
5	97	5.06	16	8	.93	23.1	−0.29	12	8	.66	32.7
7	119	3.81	12	8	.96	18.2	−2.74	16	8	.61	31.5
9	52	2.23	16	4	.96	10.0	−3.09	8	4	.81	25.0
Pe											
4	58	−18.73	16	4	.93	23.1	−7.35	–	4	.65	31.6
5	97	−17.11	8	4	.96	23.1	−4.33	8	4	.74	32.7
7	119	−15.12	4	4	.99	18.2	−0.40	8	4	.74	31.5
9	52	−13.86	12	4	.97	10.0	0.01	12	4	.75	25.0
Delta power											
4	58	−0.08	–	20	.74	23.1	0.40	28	8	.51	31.6
5	97	0.11	24	12	.88	23.1	0.33	32	8	.60	32.7
7	119	0.04	24	12	.90	18.2	0.76	12	8	.70	31.5
9	52	−0.10	28	8	.90	10.0	1.21	12	4	.71	25.0
Theta power											
4	58	−0.23	–	24	.72	23.1	0.24	16	8	.55	31.6
5	97	−0.04	–	16	.84	23.1	0.53	–	12	.35	32.7
7	119	−0.07	28	16	.90	18.2	0.75	28	8	.66	31.5
9	52	−0.27	16	8	.95	10.0	0.85	8	4	.81	25.0
Delta ITPS											
4	58	.015	–	NA	.83	23.1	.035	24	NA	.39	31.6
5	97	.016	–	NA	.88	23.1	.032	–	NA	.44	32.7
7	119	.012	–	NA	.86	18.2	.043	20	NA	.54	31.5
9	52	.007	32	NA	.93	10.0	.057	12	NA	.71	25.0
Theta ITPS											
4	58	.021	–	NA	.83	23.1	.017	–	NA	.31	31.6
5	97	.021	–	NA	.82	23.1	.024	–	NA	.31	32.7
7	119	.027	–	NA	.88	18.2	.021	28	NA	.51	31.5
9	52	.036	32	NA	.94	10.0	.016	24	NA	.57	25.0
Delta ICPS											
4	58	−.002	–	NA	.53	23.1	−.001	–	NA	.46	31.6
5	97	.001	–	NA	.64	23.1	.017	–	NA	.38	32.7
7	119	−.001	–	NA	.81	18.2	.020	28	NA	.55	31.5
9	52	−.001	–	NA	.71	10.0	.021	–	NA	.62	25.0
Theta ICPS											
4	58	−.001	–	NA	.55	23.1	.004	–	NA	.19	31.6
5	97	.000	–	NA	.53	23.1	.001	–	NA	.36	32.7
7	119	.000	–	NA	.68	18.2	.006	32	NA	.40	31.5
9	52	.006	–	NA	.74	10.0	.009	–	NA	.47	25.0

Note: Min avg = minimum number of trials needed for average reliability estimate to meet acceptable reliability (.6); Min CI = minimum number of trials needed to capture the acceptable reliability cutoff (.6) within the 95% CI; Overall ICC = overall estimate of reliability using all the available data; NA = Not applicable for this measure; "–" = acceptable reliability cutoff (.6) was not met within 32 trials.
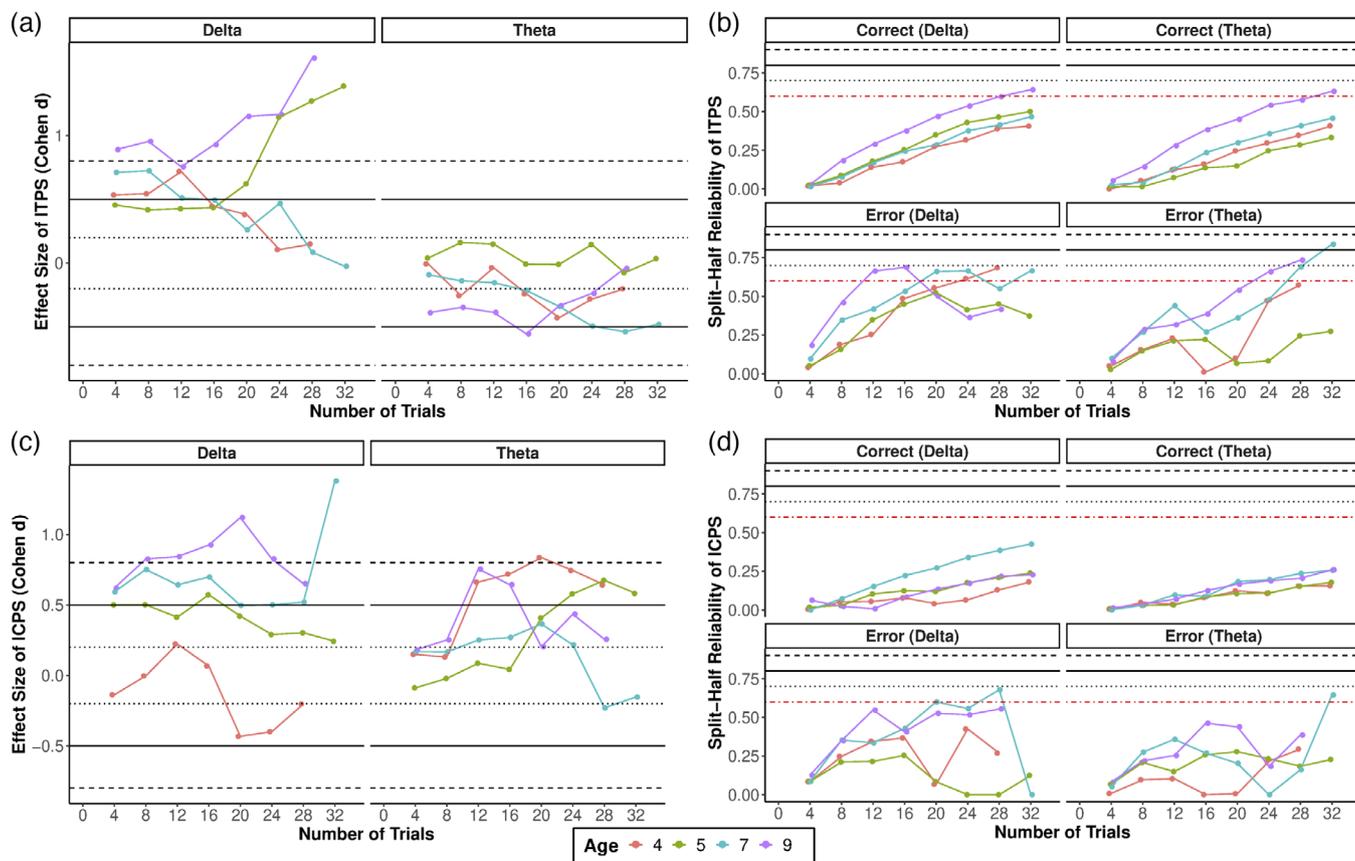
FIGURE 6 Time–frequency ITPS and ICPS effect size and reliability estimates across increasing numbers of trials for each age. Plots show the effect size of error versus correct for each ITPS (a) and ICPS (c) at each frequency, as well as split-half reliability ITPS (b) and ICPS (d) for each condition at a given frequency

between frontocentral and frontolateral areas develops across early childhood.

Together, the TF analyses provide support for the involvement of delta/theta oscillations as a neural mechanism by which the error-monitoring system signals the need for increased control and may allow for communication between structures in the error-monitoring system after errors. Our study suggests that several of these purported oscillatory neuronal mechanisms of error monitoring are present in early childhood, but also show development across childhood. Importantly, we did not observe error-specific effects when examining ERPs. This suggests that TF approaches, in addition to being more directly related to the neural mechanisms, may also be more sensitive for detecting developmental changes in error-monitoring processes.

4.3 | Psychometric analyses and effect sizes

We also examined reliability and effect size estimates of the ERP and TF measures. In order to provide valuable information for future studies planning to study error-related processes using these measures, we provide these estimates at varying numbers of trials separately at different ages. For instance, future studies may use the reliability information to help determine a minimum number of trials needed to obtain a

reliable measure or use the effect size estimates when planning a study designed to capture error-related effects at a given age. In general, ERP and TF power measures demonstrated adequate overall internal consistency reliability. However, for most measures, more trials were needed to achieve acceptable reliability than what is commonly used in the psychophysiological literature and what was used in the current study, a priori, based on previous recommendations (Meyer et al., 2014; Pontifex et al., 2010; Steele et al., 2016). TF phase-based measures demonstrated inadequate reliability, especially ICPS, suggesting that more than 32 trials are needed to reliably capture ICPS in early childhood. To our knowledge, this is the first reliability estimates of ITPS and ICPS in childhood. Given the novelty of these phase-based measures in young children, more research is needed to determine their reliability in other samples, contexts (e.g., other tasks), and ages. Overall, reliability estimates for ERPs and TF-based measures followed a similar pattern in that they increased with age. This is likely because younger ages tended to have fewer trials, smaller condition effects, possibly lower signal-to-noise ratio, and thus lower reliability for most measures.

Future studies focused on error-related processes using ERP and TF measures with young children should attempt to increase the number of error trials. This could be done by extending the length of the task or by utilizing adaptive tasks that ensure a given number of errors (e.g., Brooker, 2018). Future developmental EEG work could also

follow a multisession approach recently employed in other neuroimaging modalities (e.g., fMRI; Ellis et al., 2020) and assess the same participant repeatedly in a short period of time to ensure enough data are collected in each condition of interest. Finally, one of the most important determinants of reliability is data quality and high signal-to-noise (Clayson, 2020); thus, studies should focus on collecting high quality data and continue efforts to develop preprocessing methods that reduce the influence of artifacts and maximize the data collected.

Notably, the reliability estimates provided here should only be used to help inform future studies due to the fact that reliability estimates are a property of the scores rather than the measure. Reliability estimates are context dependent and ideally should be examined and reported in each study. For instance, reliability recommendations for the ERN range from two trials to over 40 trials, depending on the sample and paradigm with Go/No-Go tasks providing higher reliability than other tasks (Clayson, 2020). To facilitate other studies estimating the minimum number of trials needed to achieve acceptable reliability, a given effect size, or the overall reliability, we have made publicly available all the R-based functions used in the current study to estimate reliability and effect sizes (https://github.com/SMoralesPhD/SplitHalf_Reliability). Similarly, we also encourage the use of other toolboxes that perform similar reliability analyses (Clayson & Miller, 2017).

4.4 | Limitations and future directions

The current study has several limitations. First, not all participants enrolled in the study provided usable EEG data, leading to missing data. Importantly, our analyses examining factors associated with missing data indicated that missing data did not occur at random. Missing data were more likely to occur for younger participants and for participants from under-represented minorities in developmental cognitive neuroscience research (i.e., non-White participants and low levels of maternal education). Future research should continue efforts to increase diversity and representation in psychophysiological research samples by designing paradigms and research protocols that reduce participant burden and limit missing data. We attempted to make the task more developmentally appropriate for the 4-year-old group by only using one No-Go stimuli exemplar, rather than three. This change in the task is a limitation that could make age-related comparisons more difficult to interpret, as changes to other task parameters (e.g., durations of stimuli and intertrial intervals, or Go/No-Go ratios) have been shown to drive changes in behavioral performance and ERPs (Wessel, 2018; Young et al., 2018). However, in the current study, we held constant the task parameters that have previously been shown to have an influence on ERPs and behavior, and only changed the number of No-Go exemplars. Moreover, given the differential data loss associated with age, the current study is likely underestimating age-related effects in error-monitoring processes. The participants from younger ages included in the paper are likely higher in self-regulation (and error monitoring) than the participants who were excluded because of inadequate performance or not enough artifact-free trials. It is also important to note that the children in this study completed two other EEG tasks before completing the Go/No-Go task and the visit was embedded in a larger protocol involving other health (e.g., anthropometry and spirometry) and cognitive assessments. Consequently, the current study is likely a conservative estimate of the amount of error-related data that can be successfully collected in young children. Second, although the current study is one of the first to examine age-related changes in error-related processes using ERP and TF measures, we had a relatively limited age range. Moreover, it was a cross-sectional study with relatively narrow age ranges around each age 4, 5, 7, and 9 years. Future studies utilize designs that capture within-person changes across a wide age range (i.e., accelerated cohort longitudinal design) are needed to better understand the development of error monitoring. Although in the current study we show reliability differences between measures (e.g., ERN vs. Pe) and between ages (with older children having more reliable measures), future studies should examine what other factors impact the reliability of EEG measures in developmental populations such as data quality, sex, behavioral performance (e.g., RTs), or other methodological decisions (e.g., different tasks or single electrodes vs electrode clusters; Clayson, 2020). Determining what factors impact reliability will be crucial to provide EEG-based measures that can adequately differentiate between participants across development. Finally, although theta ICPS is commonly interpreted as connectivity, caution is warranted when referring to delta ICPS as connectivity because there are fewer studies focused on this frequency band; thus, we use the term synchrony instead. Further research is needed to determine if delta ICPS is implicated in neuronal communication across brain regions in a similar way as theta ICPS.

5 | CONCLUSIONS

The current study utilized ERP and TF approaches to provide a thorough examination of error-related processes across childhood (4–9 years). Across measures, the study provides evidence for the early development of the error-monitoring system. Specifically, ERP analyses showed no error-specific age-related changes in the ERN, but the Pe increased at older ages. TF analyses showed age-related changes in delta and theta signal strength (power), delta consistency (ITPS), and delta synchrony between frontrocentral (ACC) and frontolateral areas (dlPFC). Moreover, we provide reliability and effect size estimates of the ERP and TF measures across age, providing information for future studies examining the development of error monitoring. Finally, we also provide publicly available tools to facilitate and encourage each EEG study to measure their reliability and effect size estimates.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Santiago Morales https://orcid.org/0000-0002-9850-042X
Maureen E. Bowers https://orcid.org/0000-0002-4233-723X

REFERENCES

Aviyente, S., Bernat, E. M., Evans, W. S., & Sponheim, S. R. (2011). A phase synchrony measure for quantifying dynamic functional integration in the brain. *Human Brain Mapping*, *32*(1), 80–93. https://doi.org/10.1002/hbm.21000

Boldt, A., & Yeung, N. (2015). Shared neural markers of decision confidence and error detection. *Journal of Neuroscience*, *35*(8), 3478–3484. https://doi.org/10.1523/JNEUROSCI.0797-14.2015

Bowers, M. E., Buzzell, G. A., Bernat, E. M., Fox, N. A., & Barker, T. V. (2018). Time-frequency approaches to investigating changes in feedback processing during childhood and adolescence. *Psychophysiology*, *55*(10), e13208. https://doi.org/10.1111/psyp.13208

Bowers, M. E., Morales, S., Buzzell, G. A., & Fox, N. A. (2021). The influence of monetary reward on proactive and reactive control in adolescent males. *Developmental Cognitive Neuroscience*, *48*, 100934. https://doi.org/10.1016/j.dcn.2021.100934

Brooker, R. J. (2018). Maternal behavior and socioeconomic status predict longitudinal changes in error-related negativity in preschoolers. *Child Development*, *89*(3), 725–733. https://doi.org/10.1111/cdev.13066

Buzzell, G. A., Barker, T. V., Troller-Renfree, S. V., Bernat, E. M., Bowers, M. E., Morales, S., Bowman, L. C., Henderson, H. A., Pine, D. S., & Fox, N. A. (2019). Adolescent cognitive control, theta oscillations, and social observation. *Neuroimage*, *198*, 13–30. https://doi.org/10.1016/j.neuroimage.2019.04.077

Buzzell, G. A., Richards, J. E., White, L. K., Barker, T. V., Pine, D. S., & Fox, N. A. (2017). Development of the error-monitoring system from ages 9–35: Unique insight provided by MRI-constrained source localization of EEG. *Neuroimage*, *157*, 13–26. https://doi.org/10.1016/j.neuroimage.2017.05.045

Buzzell, G. A., Troller-Renfree, S. V., Wade, M., Debnath, R., Morales, S., Bowers, M. E., Zeanah, C. H., Nelson, C. A., & Fox, N. A. (2020). Adolescent cognitive control and mediofrontal theta oscillations are disrupted by neglect: Associations with transdiagnostic risk for psychopathology in a randomized controlled trial. *Developmental Cognitive Neuroscience*, *43*, 100777. https://doi.org/10.1016/j.dcn.2020.100777

Canen, M. J., & Brooker, R. J. (2017). ERN, theta power, and risk for anxiety problems in preschoolers. *Biological Psychology*, *123*, 103–110. https://doi.org/10.1016/j.biopsycho.2016.12.003

Cavanagh, J. F., Cohen, M. X., & Allen, J. J. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *Journal of Neuroscience*, *29*(1), 98–105. https://doi.org/10.1523/JNEUROSCI.4137-08.2009

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Cavanagh, J. F., & Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: Meta-analytic evidence. *Journal of Physiology-Paris*, *109*(1–3), 3–15. https://doi.org/10.1016/j.jphysparis.2014.04.003

Clayson, P. E. (2020). Moderators of the internal consistency of error-related negativity scores: A meta-analysis of internal consistency estimates. *Psychophysiology*, *57*(8), e13583. https://doi.org/10.1111/psyp.13583

Clayson, P. E., & Miller, G. A. (2017). ERP Reliability Analysis (ERA) Toolbox: An open-source toolbox for analyzing the reliability of event-related brain potentials. *International Journal of Psychophysiology*, *111*, 68–79. https://doi.org/10.1016/j.ijpsycho.2016.10.012

Cohen, M. X. (2014). *Analyzing neural time series data: Theory and practice.* MIT Press.

Crowley, M. J., van Noordt, S. J., Wu, J., Hommer, R. E., South, M., Fearon, R. M. P., & Mayes, L. C. (2014). Reward feedback processing in children and adolescents: Medial frontal theta oscillations. *Brain and Cognition*, *89*, 79–89. https://doi.org/10.1016/j.bandc.2013.11.011

Davies, P. L., Segalowitz, S. J., & Gavin, W. J. (2004). Development of response-monitoring ERPs in 7-to 25-year-olds. *Developmental Neuropsychology*, *25*(3), 355–376. https://doi.org/10.1207/s15326942dn2503_6

Debnath, R., Buzzell, G. A., Morales, S., Bowers, M. E., Leach, S. C., & Fox, N. A. (2020). The Maryland analysis of developmental EEG (MADE) pipeline. *Psychophysiology*, *57*(6), e13580. https://doi.org/10.1111/psyp.13580

Debner, S., Thorne, J., Schneider, T. T., & Viola, F. C. (2010). Using ICA for the analysis of multi-channel EEG data. In (M. Ullsperger & S. Debner Eds.), *Simultaneous EEG and fMRI* (pp. 121–135). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195372731.003.0008

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dukes, K., Burd, L., Elliott, A., Fifer, W., Folkerth, R., Hankins, G., Hereld, D., Hoffman, H., Myers, M., & Odendaal, H. (2014). The safe passage study: Design, methods, recruitment, and follow-up approach. *Paediatric and Perinatal Epidemiology*, *28*(5), 455–465. https://doi.org/10.1111/ppe.12136

DuPuis, D., Ram, N., Willner, C. J., Karalunas, S., Segalowitz, S. J., & Gatzke-Kopp, L. M. (2015). Implications of ongoing neural development for the measurement of the error-related negativity in childhood. *Developmental Science*, *18*(3), 452–468. https://doi.org/10.1111/desc.12229

Durston, S., Thomas, K. M., Yang, Y., Ulug, A. M., Zimmerman, R. D., & Casey, B. J. (2002). A neural basis for the development of inhibitory control. *Developmental Science*, *5*(4), F9–F16. https://doi.org/10.1111/1467-7687.00235

Ellis, C. T., Skalaban, L. J., Yates, T. S., Bejjanki, V. R., Córdova, N. I., & Turk-Browne, N. B. (2020). Re-imagining fMRI for awake behaving infants. *Nature Communications*, *11*(1), 4523. https://doi.org/10.1038/s41467-020-18286-y

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, *78*(6), 447–455. https://doi.org/10.1016/0013-4694(91)90062-9

Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: A tutorial. *Biological Psychology*, *51*(2–3), 87–107. https://doi.org/10.1016/S0301-0511(99)00031-9

Fox, N. A., Buzzell, G. A., Morales, S., Valadez, E. A., Wilson, M., & Henderson, H. A. (2020). Understanding the emergence of social anxiety in children with behavioral inhibition. *Biological Psychiatry*, *89*, 681-689, https://doi.org/10.1016/j.biopsych.2020.10.004

Gavin, W. J., Lin, M.-H., & Davies, P. L. (2019). Developmental trends of performance monitoring measures in 7-to 25-year-olds: Unraveling the complex nature of brain measures. *Psychophysiology*, *56*(7), e13365. https://doi.org/10.1111/psyp.13365

Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation.

Psychological Science, 4(6), 385–390. https://doi.org/10.1111/j.1467-9280.1993.tb00586.x

Gillman, M. W., & Blaisdell, C. J. (2018). Environmental Influences on Child Health Outcomes, a research program of the NIH. *Current Opinion in Pediatrics*, 30(2), 260. https://doi.org/10.1097/MOP.0000000000000600

Grammer, J. K., Carrasco, M., Gehring, W. J., & Morrison, F. J. (2014). Age-related changes in error processing in young children: A school-based investigation. *Developmental Cognitive Neuroscience*, 9, 93–105. https://doi.org/10.1016/j.dcn.2014.02.001

Grammer, J. K., Gehring, W. J., & Morrison, F. J. (2018). Associations between developmental changes in error-related brain activity and executive functions in early childhood. *Psychophysiology*, 55(3), e13040. https://doi.org/10.1111/psyp.13040

Hajcak, G., McDonald, N., & Simons, R. F. (2003). Anxiety and error-related brain activity. *Biological Psychology*, 64(1–2), 77–90. https://doi.org/10.1016/S0301-0511(03)00103-0

He, J., Degnan, K. A., McDermott, J. M., Henderson, H. A., Hane, A. A., Xu, Q., & Fox, N. A. (2010). Anger and approach motivation in infancy: Relations to early childhood inhibitory control and behavior problems. *Infancy*, 15(3), 246–269. https://doi.org/10.1111/j.1532-7078.2009.00017.x

Kikuchi, M., Shitamichi, K., Yoshimura, Y., Ueno, S., Remijn, G. B., Hirosawa, T., Munesue, T., Tsubokawa, T., Haruta, Y., Oi, M., Higashida, H., & Minabe, Y. (2011). Lateralized theta wave connectivity and language performance in 2- to 5-year-old children. *Journal of Neuroscience*, 31(42), 14984–14988. https://doi.org/10.1523/JNEUROSCI.2785-11.2011

Kim, M. H., Grammer, J. K., Marulis, L. M., Carrasco, M., Morrison, F. J., & Gehring, W. J. (2016). Early math and reading achievement are associated with the error positivity. *Developmental Cognitive Neuroscience*, 22, 18–26. https://doi.org/10.1016/j.dcn.2016.09.002

Kim, S. H., Buzzell, G., Faja, S., Choi, Y. B., Thomas, H. R., Brito, N. H., Shuffrey, L. C., Fifer, W. P., Morrison, F. D., Lord, C., & Fox, N. (2020). Neural dynamics of executive function in cognitively able kindergarteners with autism spectrum disorders as predictors of concurrent academic achievement. *Autism*, 24(3), 780–794. https://doi.org/10.1177/1362361319874920

Ladouceur, C. D., Dahl, R. E., & Carter, C. S. (2007). Development of action monitoring through adolescence into adulthood: ERP and source localization. *Developmental Science*, 10(6), 874–891. https://doi.org/10.1111/j.1467-7687.2007.00639.x

Leach, S. C., Morales, S., Bowers, M. E., Buzzell, G. A., Debnath, R., Beall, D., & Fox, N. A. (2020). Adjusting ADJUST: Optimizing the ADJUST algorithm for pediatric data using geodesic nets. *Psychophysiology*, 57, e13566. https://doi.org/10.1111/psyp.13566

Lo, S. L., Schroder, H. S., Moran, T. P., Durbin, C. E., & Moser, J. S. (2015). Neurophysiological evidence of an association between cognitive control and defensive reactivity processes in young children. *Developmental Cognitive Neuroscience*, 15, 35–47. https://doi.org/10.1016/j.dcn.2015.09.001

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT Press.

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development*, 75(5), 1357–1372. https://doi.org/10.1111/j.1467-8624.2004.00745.x

MATLAB. (2014). version 8.4 (R2014b). Natick, Massachusetts: The MathWorks Inc.

Marshall, P. J., Bar-Haim, Y., & Fox, N. A. (2002). Development of the EEG from 5 months to 4 years of age. *Clinical Neurophysiology*, 113(8), 1199–1208. https://doi.org/10.1016/S1388-2457(02)00163-3

Meyer, A., Bress, J. N., & Proudfit, G. H. (2014). Psychometric properties of the error-related negativity in children and adolescents: Psychometric properties of ERN in youth. *Psychophysiology*, 51(7), 602–610. https://doi.org/10.1111/psyp.12208

Morales, S., & Bowers, M. (2021). *Time-frequency analysis methods and their application in developmental EEG data*. PsyArXiv. https://doi.org/10.31234/osf.io/dcek3

Morales, S., Miller, N. V., Troller-Renfree, S. V., White, L. K., Degnan, K. A., Henderson, H. A., & Fox, N. A. (2020). Attention bias to reward predicts behavioral problems and moderates early risk to externalizing and attention problems. *Development and Psychopathology*, 32(2), 397–409. https://doi.org/10.1017/S0954579419000166

Morales, S., Taber-Thomas, B. C., & Pérez-Edgar, K. E. (2016). Patterns of attention to threat across tasks in behaviorally inhibited children at risk for anxiety. *Developmental Science*, 20(2), e12391. https://doi.org/10.1111/desc.12391

Munneke, G.-J., Nap, T. S., Schippers, E. E., & Cohen, M. X. (2015). A statistical comparison of EEG time- and time–frequency domain representations of error processing. *Brain Research*, 1618, 222–230. https://doi.org/10.1016/j.brainres.2015.05.030

Narayanan, N. S., Cavanagh, J. F., Frank, M. J., & Laubach, M. (2013). Common medial frontal mechanisms of adaptive control in humans and rodents. *Nature Neuroscience*, 16(12), 1888–1895. https://doi.org/10.1038/nn.3549

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P. H., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology*, 38(5), 752–760. https://doi.org/10.1111/1469-8986.3850752

Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: Fully automated statistical EEG artifact rejection. *Journal of Neuroscience Methods*, 192(1), 152–162. https://doi.org/10.1016/j.jneumeth.2010.07.015

Orekhova, E. V., Stroganova, T. A., Posikera, I. N., & Elam, M. (2006). EEG theta rhythm in infants and preschool children. *Clinical Neurophysiology*, 117(5), 1047–1062. https://doi.org/10.1016/j.clinph.2005.12.027

Overbeek, T. J., Nieuwenhuis, S., & Ridderinkhof, K. R. (2005). Dissociable components of error processing: On the functional significance of the Pe vis-à-vis the ERN/Ne. *Journal of Psychophysiology*, 19(4), 319–329. https://doi.org/10.1027/0269-8803.19.4.319

Pasion, R., & Barbosa, F. (2019). ERN as a transdiagnostic marker of the internalizing-externalizing spectrum: A dissociable meta-analytic effect. *Neuroscience & Biobehavioral Reviews*, 103, 133–149. https://doi.org/10.1016/j.neubiorev.2019.06.013

Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D., & R Core Team. (2016). *nlme: Linear and Nonlinear Mixed Effects Models* (3.1-124) [Computer software]. http://CRAN.R-project.org/package=nlme

Pontifex, M. B., Scudder, M. R., Brown, M. L., O'Leary, K. C., Wu, C.-T., Themanson, J. R., & Hillman, C. H. (2010). On the number of trials necessary for stabilization of error-related brain activity across the life span. *Psychophysiology*, 47(4), 767–773.

R Development Core Team. (2008). *R: A language and environment for statistical computing*. http://www.R-project.org

Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews Neuroscience*, 12(3), 154–167. https://doi.org/10.1038/nrn2994

Steele, V. R., Anderson, N. E., Claus, E. D., Bernat, E. M., Rao, V., Assaf, M., Pearlson, G. D., Calhoun, V. D., & Kiehl, K. A. (2016). Neuroimaging measures of error-processing: Extracting reliable signals from event-related potentials and functional magnetic resonance imaging. *Neuroimage*, 132, 247–260. https://doi.org/10.1016/j.neuroimage.2016.02.046

Steinhauser, M., & Yeung, N. (2010). Decision processes in human performance monitoring. *Journal of Neuroscience*, 30(46), 15643–15653. https://doi.org/10.1523/JNEUROSCI.1899-10.2010

Tamnes, C. K., Walhovd, K. B., Torstveit, M., Sells, V. T., & Fjell, A. M. (2013). Performance monitoring in children and adolescents: A review of developmental changes in the error-related negativity and brain maturation. *Developmental Cognitive Neuroscience*, 6, 1–13. https://doi.org/10.1016/j.dcn.2013.05.001

Tenke, C. E., & Kayser, J. (2012). Generator localization by current source density (CSD): Implications of volume conduction and field closure

at intracranial and scalp resolutions. *Clinical Neurophysiology*, *123*(12), 2328–2345. https://doi.org/10.1016/j.clinph.2012.06.005

Thorpe, S. G., Cannon, E. N., & Fox, N. A. (2016). Spectral and source structural development of mu and alpha rhythms from infancy through adulthood. *Clinical Neurophysiology*, *127*(1), 254–269. https://doi.org/10.1016/j.clinph.2015.03.004

Torpey, D. C., Hajcak, G., Kim, J., Kujawa, A., & Klein, D. N. (2012). Electrocortical and behavioral measures of response monitoring in young children during a Go/No-Go task. *Developmental Psychobiology*, *54*(2), 139–150. https://doi.org/10.1002/dev.20590

Towers, D. N., & Allen, J. J. (2009). A better estimate of the internal consistency reliability of frontal EEG asymmetry scores. *Psychophysiology*, *46*(1), 132–142. https://doi.org/10.1111/j.1469-8986.2008.00759.x

Troller-Renfree, S. V., Buzzell, G. A., Bowers, M. E., Salo, V. C., Forman-Alberti, A., Smith, E., Papp, L. J., McDermott, J. M., Pine, D. S., Henderson, H. A., & Fox, N. A. (2019). Development of inhibitory control during childhood and its relations to early temperament and later social anxiety: Unique insights provided by latent growth modeling and signal detection theory. *Journal of Child Psychology and Psychiatry*, *60*(6), 622–629. https://doi.org/10.1111/jcpp.13025

Wessel, J. R. (2018). Prepotent motor activity and inhibitory control demands in different variants of the go/no-go paradigm. *Psychophysiology*, *55*(3), e12871. https://doi.org/10.1111/psyp.12871

Young, M. E., Sutherland, S. C., & McCoy, A. W. (2018). Optimal go/no-go ratios to maximize false alarms. *Behavior Research Methods*, *50*(3), 1020–1029. https://doi.org/10.3758/s13428-017-0923-5

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). II. NIH Toolbox Cognition Battery (CB): Measuring executive function and attention. *Monographs of the Society for Research in Child Development*, *78*(4), 16–33. https://doi.org/10.1111/mono.12032

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.