Contents lists available at ScienceDirect



Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev



Review article

A meta-analytic evaluation of the N2 component as an endophenotype of response inhibition and externalizing psychopathology in childhood



Caroline P. Hoyniak^{a,*}, Isaac T. Petersen^b

^a Indiana University, Department of Psychological and Brain Sciences, 1101 E. 10thStreet A304, Bloomington, IN, 46220, United States ^b University of Iowa, Department of Psychological and Brain Sciences, 308 Iowa Avenue, SLP Room 111, Iowa City, IA, 52242, United States

ARTICLE INFO

Keywords: Event-related potential N200 Response inhibition Self-regulation Externalizing psychopathology

ABSTRACT

The N2 event-related potential component is a well-studied neurophysiological index of response inhibition that is considered to be a biomarker of externalizing psychopathology. The literature on the N2 elicited in childhood has been inconsistent, though, with different studies yielding different findings regarding the association between the N2 and the constructs it is thought to index. The current meta-analysis sought to clarify the functional meaning of the N2 component elicited in childhood across three widely used response inhibition tasks. The current study meta-analyzed the findings of 54 studies examining the association of the N2 component and three phenotypes of interest: (1) behavioral response inhibition (as indexed by performance on the inhibition trials of the task used to elicit the N2 component), (2) performance on behavioral measures of self-regulation, and (3) psychopathology (both externalizing and internalizing) in samples of children, to clarify the meaning of the N2 component is associated with response inhibition and externalizing psychopathology.

1. Introduction

There is growing recognition that the diagnostic categories described in the Diagnostic and Statistical Manual of Mental Disorders (DSM) are heterogeneous, and that a single set of symptoms (e.g., anhedonia) can appear across a multitude of diagnostic categories (e.g., depression and schizophrenia). The heterogeneity of the diagnostic categories and their extensive comorbidity likely, in part, reflects that the DSM defines mental disorders based on their phenotypic expressions (e.g., behaviors). Principles in developmental psychopathology, such as equifinality, indicate that different etiologies can lead to the same outcome behavior. Thus, downstream and heterogeneous symptom-based definitions of psychopathology may not carve nature at its joints. Instead, understanding underlying neurodevelopmental trajectories of psychopathology is a critical need that will advance our understanding of psychopathology, how it develops, and how to treat it (Casey et al., 2014; Insel, 2014).

The desire for researchers to move away from a focus on the symptom-based diagnostic categories outlined in the DSM is especially apparent in the National Institute of Mental Health's Research Domain Criteria (RDoC) initiative. Within the RDoC framework, researchers are encouraged to identify the biological underpinnings of domains of psychopathology, in order to map brain-behavior, and eventually genome-behavior, associations (Insel et al., 2010). There is a focus on studying psychopathology from multiple levels of analysis, integrating findings from cellular and genetic studies, to what is known about brain networks and psychophysiology, and eventually to more readily visible behavior and symptom clusters.

The current study seeks to advance our understanding of neural processes in the development of externalizing psychopathology in childhood from a multilevel, RDoC perspective. Externalizing behavior problems, which consist of acting-out behaviors such as aggression, inattention, hyperactivity, and conduct problems, are frequent, costly, and burdensome in children. The worldwide prevalence of externalizing disorders (e.g., conduct disorder [CD] and oppositional defiant disorder [ODD]) in childhood is estimated to be 5.7%, or 113 million children (Polanczyk et al., 2015). Additionally, individual differences in externalizing problems are highly stable from childhood to adulthood, with externalizing problems in childhood predicting severe outcomes in adulthood including substance use and criminality (Petersen et al., 2015). Thus, it is crucial to understand etiological processes in the development of externalizing problems in childhood.

The N2 event-related potential (ERP) is among the most widely studied neural markers of externalizing psychopathology. The N2 is thought to be an index of response inhibition (discussed in more detail below, Section 1.1), and is thought to be a biomarker of externalizing

* Corresponding author.

E-mail addresses: choyniak@indiana.edu (C.P. Hoyniak), isaac-t-petersen@uiowa.edu (I.T. Petersen).

https://doi.org/10.1016/j.neubiorev.2019.06.011

Received 26 September 2018; Received in revised form 6 June 2019; Accepted 11 June 2019 Available online 12 June 2019

0149-7634/ © 2019 Elsevier Ltd. All rights reserved.

disorders characterized by disinhibition (e.g., externalizing psychopathology such as CD, ODD, and attention deficit hyperactivity disorder [ADHD]). We integrate across multiple levels of analysis, meta-analyzing available data to examine the associations between the N2 and behavioral expressions of disinhibition, and externalizing psychopathology symptom clusters.

1.1. The N2 ERP component

In the last few decades, a growing literature has focused on identifying the neural biomarkers of emerging psychopathology in childhood. This research has utilized a variety of methodologies, including electroencephalography (EEG) and corresponding ERPs, functional magnetic resonance imaging (fMRI), and near-infrared spectroscopy (NIRS). Because of their relatively low cost, their utility for use with young children, and their high temporal resolution, EEG/ERPs have been a popular tool for use in this area of research. Research utilizing ERPs has explored the neural correlates of a variety of forms of psychopathology, including depression (e.g., Kujawa et al., 2012), anxiety disorders (e.g., Meyer et al., 2015), callous and unemotional traits (e.g., Hoyniak et al., 2018), attention problems (e.g., Johnstone and Clarke, 2009), and disruptive behavior problems (e.g., Grabell et al., 2017), focusing on a variety of ERP components, including the N170, the errorrelated negativity (perhaps more correctly described as the reward-related positivity; Proudfit, 2015), the N2, and the P3.

The N2 ERP component, a component thought to index inhibitory capacities in tasks requiring response inhibition, has been widely studied in relation to both externalizing and internalizing psychopathology, as disinhibition is thought to be a transdiagnostic marker of psychopathology in childhood (Bunford et al., 2015). However, disinhibition has generally shown stronger associations with externalizing psychopathology than with internalizing psychopathology (Bates et al., 2014). The N2 ERP component, the second negative deflection in the waveform that occurs approximately 200-400 ms post-stimulus in children (Hoyniak, 2017) to both activation (Go) and inhibition (NoGo/ Stop) stimuli, is thought to specifically index response inhibition capacities. Response inhibition is the suppression of a behavioral response that is cued by the presentation of an inhibition signal that is either internal or external. The N2 component is noticeably larger to inhibition stimuli than it is to activation stimuli, contributing to this component's functional interpretation as an index of response inhibition. Of note, the N2 has also been studied as a "novelty" component, thought to index attention to mismatched visual stimuli (Folstein and Van Petten, 2008). However, when the N2 is elicited in the context of tasks requiring response inhibition, the functional interpretation of the component centers on cognitive control and inhibitory capacities. This N2, termed the "control" N2 (but henceforth referred to simply as the N2), is the focus of the current study. By contrast, a number of researchers have proposed that it is not inhibition that the N2 indexes, but rather response conflict (Groom and Cragg, 2015). Response conflict emerges when there is competition among various incompatible response pathways (e.g., an activation response vs. an inhibition response). As response inhibition tasks elicit response conflict, along with inhibition, some researchers argue that it is response conflict that the N2 indexes (Nieuwenhuis et al., 2003; Huster et al., 2013). There has been support for both response inhibition and response conflict interpretations of the N2, suggesting perhaps that the N2, like many other tasks and neural markers associated with higher-order cognitive processes, has a number of related functional interpretations.

In a discussion of inhibition/conflict related ERP components, we would be remiss not to mention the NoGo P300, also known as the frontal or inhibitory P300. The NoGo P300 peaks 300–500 ms poststimulus in adults, is maximal over frontocentral electrodes, and, like the NoGo N2, has been interpreted to be an index of response inhibition (Bokura et al., 2001; Lewis et al., 2006). The NoGo P300 is often examined in conjunction with the NoGo N2 (although it is not always elicited in tasks that elicit the N2 among young children; Grabell et al., 2017; Isbell et al., 2019). Although we acknowledge the importance of the NoGo P300 and encourage similar future meta-analytic investigations of the NoGo P300 explicitly, the already large scope of the current examination of the N2 prohibited us from considering this component further in the current study.

The N2 component has been elicited across various tasks assessing response inhibition, including the Go/NoGo (GNG) task, the Continuous Performance task (CPT), and the Stop-Signal Task (SST). Task choice is often influenced by a number of factors including research traditions and type of psychopathology of interest. Although the N2 component can be elicited during other tasks (e.g., the Flanker task), the following meta-analysis focuses specifically on the GNG task, the CPT, and the SST, because these are the tasks most commonly used when assessing the N2 associated with response inhibition.

Although these three tasks each require response inhibition, there are likely subtle differences in the neural bases of inhibition assessed based on task parameters and individual difference factors. For example, there may be both intra-individual and inter-individual differences in the modes of cognitive control utilized during task performance, which is associated with the neural regions recruited during task performance (Braver, 2012). The dual mechanisms of control framework (Braver, 2012) posits that there are two, non-independent modes of cognitive control, proactive control (in which goal-related information is maintained throughout the duration of a task) and reactive control (in which goal-related information is activated when it is necessitated, e.g., during interference). Differences in the activation of the two modes of cognitive control depends on a number of factors, including task specifications (e.g., the extent of interference present in a task, see Gonthier et al., 2016), participant age (Chatham et al., 2009), and individual differences (e.g., reward and threat sensitivity, impulsivity; Braver, 2012). The two modes of cognitive control have been shown to be associated with activation in different neural regions (i.e., proactive control relies on sustained activity in the lateral prefrontal, while reactive control relies on more transient activity across frontoparietal regions). As such, the mode of cognitive control activated during a given task may modulate neural activity measured by EEG. Additionally, the relative frequency with which NoGo and Go trials are presented in the task can enhance or diminish the size of the N2, theoretically by modulating an individual's prepotency to respond by increasing or decreasing the amount of response conflict associated with response stimuli. Low frequency responses that occur in the context of a high frequency, non-response stimulus, as is often the case in GNG tasks, are thought to produce increased response conflict, because the required response must compete with the bias to react in a way congruent with the prepotent, high frequency stimulus (Nieuwenhuis et al., 2003). As such, the NoGo N2 has been found to be largest in tasks where the Go trials are more probable (i.e., 80% Go trials and 20% NoGo trials), slightly diminished in tasks where Go and NoGo trials are presented equiprobably, and reversed in polarity with Go trials in tasks where Go trials are more probable (i.e., 20% Go trials and 80% NoGo trials; Nieuwenhuis et al., 2003). Finally, research with humans and animals has suggested distinct neural generators for tasks that require action restraint (i.e., inhibition of a motor behavior that occurs prior to the behavior being initiated, such as in the GNG task or the CPT) and tasks that require action cancellation (i.e., inhibition of a motor behavior that occurs after that motor behavior has already been initiated, such as in the SST; Eagle et al., 2008; Schachar et al., 2007; Swick et al., 2011). While N2 components elicited across these tasks are often considered to be directly comparable, caution is clearly warranted given the factors that have been shown to modulate the N2 (highlighted above). As such, the current study considers the N2s generated in the context of these three tasks both in combination and separately, and explicitly considers how the relative probability of Go to NoGo trials modulates resulting N2s.

1.1.1. Go/NoGo (GNG) task

In the GNG task, two stimuli are introduced to the participants: a Go stimulus, which is paired with some form of response activation, e.g., a button press, counting, and a NoGo stimulus, which is paired with motor inhibition. To establish a prepotent tendency to respond thereby making the inhibition task more difficult, the Go stimuli are often presented more frequently than the NoGo stimuli-either through initially presenting a trial block that contains only Go stimuli or through presenting Go stimuli more frequently than NoGo stimuli during trial blocks. Although it is difficult to link activation in specific neural regions to observed ERP components, techniques have been developed that allow localization of the source of the components. Studies that have used source localization techniques to identify the neural generators of the N2 component elicited during a GNG task have suggested that the N2 can be localized to the anterior cingulate cortex (Bokura et al., 2001), the orbitofrontal cortex (Bokura et al., 2001), the ventral prefrontal cortex (Lavric et al., 2004), and the dorsolateral prefrontal cortex (Lavric et al., 2004). Each of these prefrontal regions has been hypothesized to play a crucial role in supporting response inhibition (Aron, 2007; Simmonds et al., 2008; Steele et al., 2013).

1.1.2. Continuous performance task (CPT)

In the CPT, participants view a continuous stream of letters or numbers, watching for a pre-specified cue stimulus. Although the cue stimulus does not elicit a response directly, the letter or number immediately following the cue stimulus indicates whether the participant should make a motor response. For example, if the cue stimulus is an "A" and the Go stimulus is an "X", participants would only execute a Go response when the cue stimulus, "A", is followed by an "X". If any other letter follows the cue stimulus, response inhibition is required. Because participants must remember a more complicated rule pattern, and sustain attention during the continuous stream of letters, the CPT is thought to be more complex than the GNG task. Source localization studies have identified several potential neural generators of the N2 component elicited in the context of the CPT, including regions of the medial frontal cortex (near the anterior cingulate cortex; Bekker et al., 2005; Jonkman et al., 2007), and, for children, occipito-temporal and parietal regions (Jonkman et al., 2007).

1.1.3. Stop-signal task (SST)

In the SST, two types of stimuli are introduced to participants: (1) one or more Go stimuli that elicit some form of response activation (e.g., a button press), and (2) a stop signal that is presented after Go stimuli on select trials, which signals the need for motor inhibition. In the SST, unlike the GNG, the stop-signal is presented *after* the onset of the Go stimulus. As a result, performance on the SST is conceptualized as reflecting action cancellation rather than the action restraint that is typically thought to be assessed in the GNG (Swick et al., 2011). The distribution of reaction times from Go trials and unsuccessful NoGo trials can be used to estimate a stop signal reaction time (SSRT) value, which is thought to reflect the amount of time, post-stimulus, needed for the individual to successfully inhibit an initiated response. Evidence suggests that the neural generators of the stop N2 in adults include the basal ganglia, anterior midcingulate cortex, and pre-supplementary motor area (Huster et al., 2011).

1.2. The meaning of the N2

Most ERP components, the N2 included, are thought to index a multitude of cognitive processes and rely on a distributed network of neural generators. As such, it is reasonable that there is disagreement among researchers as to the functional meaning and significance of any one ERP component. Although multiple interpretations of the N2 have been proposed, perhaps the most robust interpretation has been that the N2 is an index of some form of response inhibition (it is worth mentioning, though, that the N2 can be elicited in tasks that include no

overt response inhibition, demonstrating that the N2 may index additional cognitive processes; Burle et al., 2004). The interpretation of the N2 as indexing response inhibition has been supported experimentally, through manipulating task parameters and examining the effect on the N2 (e.g., Groom and Cragg, 2015; Nieuwenhuis et al., 2004), as well as non-experimentally, through examining the association between the N2 and other outcomes (e.g., Falkenstein, Hoorman, & Hohnsbein, 1999). Examinations of the association between the N2 and response inhibition in childhood, however, have been notably more variable. Given that response inhibition is considered to be a facet of self-regulation (the physiological, attentional, cognitive, emotional, and behavioral regulatory processes that promote adaptive or goal-directed behavior: Berger, 2011: Zhou et al., 2011), and has been shown to play a key role in the development of childhood externalizing problems (Schoemaker et al., 2013; Slaats-Willemse et al., 2003; Young et al., 2009), it is possible that the N2 might also be an index of these broader constructs. To evaluate this possibility, the current study examined the association between the N2 and these three behavioral phenotypes of interest: (1) behavioral response inhibition as quantified by performance on the same behavioral task used to elicit the N2, (2) trait level self-regulation as quantified by parent reports and behavioral assessments of general self-regulatory abilities, and (3) externalizing psychopathology as quantified either by diagnostic categorization or symptom checklists. Of note, we also examined the N2 in relation to internalizing psychopathology as a test of discriminant validity of the N2.

1.3. The current study

The collection standards of and interpretations for ERP components elicited from adults are often applied down to ERPs collected from children. However, both of these practices ought to be called into question. Just as ERP collection standards should be adjusted to be developmentally appropriate (Brooker et al., 2019; DeBoer et al., 2007; Thierry, 2005), so too should functional interpretations of ERP components be adjusted to consider possible differences in the functional interpretation of ERP components across development. The presence of an association between an ERP and an external correlate in adults may be used as a guiding principle for developmental research, but research with children that does not replicate the findings of studies with adults may still be meaningful. To get a clear picture of the changes in ERPs that occur across development, this research should be approached in a theoretically informed data-driven manner, in which associations already established in adults should be re-established in children.

By accruing data from the existing studies examining the N2 component in childhood, we sought to clarify the association between the childhood N2 and the three behavioral phenotypes of interest: (1) behavioral response inhibition, (2) trait self-regulation, and (3) externalizing psychopathology. The evidence for these associations in child samples is notably inconsistent, in terms of both direction of effects and statistical significance. Through accruing across the existing literature that examines the N2 and its external correlates in child samples, we may be able to get a clearer picture of the N2 in childhood. We expected, based on general trends in the adult and child literature, that smaller N2 amplitudes during task inhibition conditions would be associated with: (1) poorer behavioral response inhibition as quantified by task performance (e.g., Falkenstein et al., 1999; Fogarty et al., 2018; Van Boxtel et al., 2001), (2) poorer trait self-regulation (e.g., Ruberry et al., 2017), and (3) increased externalizing psychopathology (e.g., Woltering, Li, Rokeach, & Tannock, 2013). Smaller N2 amplitudes may reflect diminished activation of the neural regions associated with the N2 component. If the N2 is indeed associated with response inhibition, this could reflect less activation in the regions of the brain associated with response inhibition, and thus poorer response inhibition capacities. This could, in turn, feed into poorer trait self-regulation and increased externalizing problems. To our knowledge, this is the first metaanalysis to examine the functional meaning of the N2 component in

childhood in terms of response inhibition and self-regulation, and to examine the association between the N2 and externalizing problems.

2. Methods

The current meta-analysis adheres to the reporting guidelines outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Liberati et al., 2009; Moher et al., 2009) statement. Several additional sources were consulted for guidance regarding meta-analytic procedures and analyses (Cooper, 2009; Cooper et al., 2009; Lipsey and Wilson, 2001; Schwarzer et al., 2015).

2.1. Literature search

2.1.1. Study selection

One electronic database, Google Scholar, was used to generate a list of studies for the initial literature review using the following search terms in various combinations (depending on which task, GNG, CPT, or Stop Signal, was the focus of the search): N2, N200, event-related potential, ERP, child*, NoGo, No-Go, Go, stop task, stop-signal, continuous performance task, and CPT. No restrictions regarding date of publication, reference type, or author's country of origin were placed on the initial literature search, but only studies published in English were reviewed. For the current study, we added to the literature search conducted for a previous meta-analysis focusing on the GNG task and the CPT in which 1243 studies were identified in the summer of 2015 (Hoyniak, 2017). This previous meta-analysis, although also focused on the N2 component, had a vastly different scope and aims. Hoyniak (2017) focused on describing the developmental trends of the N2 component across childhood, with a particular focus on examining if there was a difference between Go and NoGo N2 amplitudes in childhood. The current study, which examines the functional meaning (i.e., criterion validity) of the N2 component in association with the behavioral constructs it is thought to index, has a different scope, examining entirely different variables (e.g., response inhibition, externalizing problems, and internalizing problems). For the current study, the literature search from Hoyniak (2017) was updated and extended in the summer of 2017, which included the identification of 398 additional studies published on the GNG task and the CPT since 2015. Additionally, the current study also examined the N2 elicited in the context of the SST (a task that Hoyniak [, 2017] did not examine), and 1040 studies published on the stop signal task were identified. Overall, 2681 studies were identified in the initial literature search.

The titles and abstracts of these studies were hand-screened by the authors to assess eligibility for inclusion. Studies that reported measuring the N2 ERP component in the context of a GNG task, CPT, or SST in a developmental sample (defined here as up to 12 years) were selected for further review (N = 167). The cutoff age of 12 was selected because it corresponds generally to the typical onset of puberty and adolescence. The authors then reviewed the full text articles for each study meeting inclusion criteria (N = 167) to determine if the study should be *excluded* based on the following pre-specified exclusion criteria:

- 1 The study included a clinical sample that was neither an externalizing sample (i.e., behavior problems or ADHD), an internalizing sample (i.e., depression or anxiety), nor a comorbid externalizing and internalizing sample. Examples of excluded clinical samples include children with Autism Spectrum Disorders, Specific Language Impairment, Fetal Alcohol Syndrome, Tic Disorders, and children born premature. While such samples were excluded, control samples or externalizing/internalizing samples from the study could be included if these samples were described and analyzed separately.
- 2 The study's sample was an exact duplicate of another included study. In such a situation, the more recently published study, or the

study in which more information was provided, was included.

- 3 The study used a variant of either the GNG task, the CPT, or the SST that did not adhere to our operational definition of each task. A GNG task was defined as a task in which at least two types of stimuli were displayed, at least one "Go" stimulus that elicited response activation (e.g. button press, counting, etc), and at least one "NoGo" stimulus that elicited response inhibition. A CPT was defined as a task that included at least one pre-specified cue stimulus that was presented among a continuously appearing stream of distractor stimuli. The cue stimulus signaled a potential need for response activation, but, depending on the letter following the cue stimulus, could also signal a need for response inhibition. A SST was defined as a task in which at least one "Go" stimulus, eliciting response activation, was presented at every trial, while on a subset of trials, a "Stop-Signal," eliciting action cancellation, was presented after some delay following the onset of the "Go" stimulus. Across all three task types, the number of presented inhibition (NoGo/Stop) stimuli must have been equal to or less than the number of presented activation (Go) stimuli.
- 4 The study did not report examining the N2 ERP component.
- 5 The study included some type of intervention, but did not provide information about ERPs gathered from either the control group (i.e., non-intervention group) or ERPs measured in the intervention group prior to the intervention. Of note, from these intervention studies, data from pre-intervention time points and control groups were included in the meta-analysis.
- 6 The study did not report an association with either task performance, trait self-regulation, or externalizing or internalizing psychopathology.

Of the 167 studies screened, 113 were excluded based on the exclusion criteria, leaving 54 studies to be included in the meta-analysis. See Fig. 1 for an overview of the literature search procedure, including information on the number of studies excluded based on each set of exclusion criteria.

Most of the 54 included studies contained multiple, unique subsamples that divided up the full sample. Subsamples could include groupings by age (e.g., 7-year-olds vs. 8-year-olds), or by diagnostic category (e.g., children with ADHD vs. controls), or by behavioral phenotype being examined (i.e., an analysis of the association between the N2 and behavioral response inhibition and the N2 and externalizing problems in the same study would be counted as two separate subsamples because they would each inform separate effect size analyses). Based on our criteria, 83 independent subsamples were extracted from the 54 included studies. Of the 83 subsamples, 34 were included in the behavioral response inhibition analysis, 6 were included in the trait self-regulation analysis, and 43 were included in the psychopathology analysis. Basic information about each study included in analysis, and the outcome variables they examined, is provided in Table 1, and citations for these studies are provided in Supplementary Appendix S1.

2.1.2. Data extraction: missing information

For each subsample, the authors extracted the following key pieces of information: (1) the number of children included in ERP analysis, (2) age of participants (range or M and SD), (3) the association between the N2 amplitude and an index of performance on the inhibitory trials of the ERP task (e.g., NoGo percent correct), (4) the association between the N2 amplitude and a broader index of self-regulation, and (5) the association between the N2 amplitude and psychopathology (e.g., group differences in the N2 between children with and without ADHD). If any of these key pieces of information were calculated, but not reported, primary authors were contacted in the summer of 2017 to give them the opportunity to provide us with the missing information. The authors of 24 studies were contacted, and 7 responded, with 4 providing the requested information.

A notable difficulty with this literature is that authors often report



Fig. 1. Overview of literature search method.

an effect size to be non-significant, but do not provide the requisite information (e.g., means or SDs) for this effect size to be included in meta-analytic combination. In the current meta-analysis, the effect sizes for 11 subsamples were reported to be non-significant, and no additional descriptive information was provided about these effect sizes nor did the corresponding authors provide us with this requested information. In this case, two approaches, both with notable limitations, can be taken. First, the meta-analysis could be limited to only studies reporting sufficient effect size information; however, this may upwardly bias the estimate if non-significant findings are systematically excluded (Lipsey and Wilson, 2001). Second, the effect size of non-significant findings could be imputed to be zero; however, this may downwardly bias estimates (Lipsey and Wilson, 2001). Without information about the direction of the effect size, we opted to pursue both approaches, noting the limitations of both, to try to understand how our metaanalytic estimates were affected by the imputation of an effect size of 0 in these 11 subsamples.

Similarly, it has been well established that study publication is a non-random event, with significant findings much more likely to be published than non-significant findings (Dickersin, 2005). Meta-analyses that only include published studies are likely to result in skewed effect size estimates, as significant studies have a disproportionate influence on the overall effect size calculations (Sutton, 2009). This may be especially relevant to neuroscience research on children, as researchers who do not replicate ERP findings with adults may have a difficult time publishing their research. Although publication bias is unavoidable, the current study used two approaches to minimize the effect of publication bias on effect size estimates. First, when possible, unpublished studies (including theses and dissertations) were included in analysis. In the current study, three publically available dissertations were included in analysis. However, even the inclusion of these types of sources does not fully negate the effect of publication bias, so we also examined if statistical correction for bias was necessary for each combined effect size estimate, described below.

2.1.3. Data extraction: predictor and outcome variables

Based on previous research, two aspects of the N2 component are thought to specifically index response inhibition capacities: (1) the inhibition (NoGo/Stop) N2 amplitude and (2) the N2 difference waveform. Only effects that examined the association between either the inhibition N2 amplitude or the N2 difference waveform were extracted from eligible subsamples. We did not consider N2 latencies, because much less research has focused on the association of N2 latencies and any of the behavioral phenotypes examined in the current study. To include the greatest number of studies possible, few restrictions were placed on the types of outcomes that could be included in each phenotype category. For the behavioral response inhibition category, outcomes included commission errors or false alarm rate (i.e., displaying behavioral activation during an inhibition trial), percent correct on NoGo trials, percent correct on the task as a whole, and accuracy-cost calculations (i.e., composites that take into account performance on both Go and NoGo trials) for the GNG task and CPT, and performance on failed inhibition trials vs. successful inhibition trials, percentage of successful inhibition trials, SSRT, and inhibition slope (i.e., the percent of failed inhibitions as a function of Go-Stop delay interval) for the Stop-Signal task. For trait self-regulation, outcomes included performance on tasks assessing self-regulation, including parent impressions of the child's regulatory skills, the Attention Network Task (ANT; Fan et al., 2002), the Stroop task (Stroop, 1935), the Bear/Dragon task (Reed et al., 1984), the Grass/Snow task (Carlson and Moses, 2001), the Day/Night task (Gerstadt et al., 1994), the Dimensional Change Card Sort task (Frye et al., 1995), the Iowa Gambling Task (Bechara et al., 1994), and various tasks assessing delay discounting abilities. As described above, the types of psychopathology outcomes eligible for inclusion were externalizing, internalizing, or comorbid externalizing and internalizing problems. Psychopathology could be defined by diagnoses using DSM-based criteria or scores on well-established symptom checklists.

Frequently, multiple outcome variables were included as indexes of the same behavioral phenotype within a single study (e.g., a study that includes both a Stroop task and an ANT as indexes of self-regulation). This is no problem if different samples are used in these different associations (e.g., different groups of children participate in the Stroop task and the ANT). However, it is more common for the same sample to be examined across multiple outcomes (e.g., the same group of children participates in both the Stroop task and the ANT). Including all of these associations as separate effect sizes in meta-analytic combination violates the assumption of independent samples (Lipsey and Wilson, 2001). Prior to meta-analytic combination, researchers need to create an independent effect size dataset (Lipsey and Wilson, 2001). In the current study, if multiple effect sizes were extracted for the same behavioral phenotype within a single subject group, these multiple effect sizes were averaged together. The average was used in meta-analytic combination analysis. This approach is consistent with best practices and is often advocated over the use of nested approaches (i.e., effect size combinations that account for multiple effect sizes; Lipsey and Wilson, 2001). A data dictionary outlining the variables extracted from each study is provided on Open Science Framework (https://osf.io/t6hwg/).

Table 1 Descriptive information ab	out th	e studies meeting in	clusion criteria.				
Study	z	Age Range (years)*	Predictor	Task	Outcomes		
					Neuropsychological Inhibition	Trait Self-Regulation	Psychopathology
Barac et al. (2016)	62	4.42-6.33	NoGo N2 Amp	GNG	Discriminability Index (d ¹)	1	1
Barry and De Blasio (2015) Benikos and Johnstone	40 36	8–13 7.75–14.6	NoGo N2 Amp NoGo N2 Amp	BNG DND	Commission Errors -	1 1	- ADHD
(2009) [2009]	Ļ	01100					
Bluschke et al. (2016) Bluschke et al. (2017)	88	9.2-15.9 9.8-17.3	NoGo N2 Amp NoGo N2 Amp	טאט טאט	Commission Errors	1 1	ADHD ADHD
Boucher et al. (2014)	186	9.8-12.9	NoGo N2 Amp	GNG	Percent Correct (NoGo)	1	ADHD: Ext Symptoms
Chevalier et al. (2014)	30	5-6	NoGo N2 Amp	GNG	Percent Correct (NoGo)	1	
Cragg et al. (2009)	56	7.28–9.76	NoGo N2 Amp	GNG	Percent Correct (NoGo)	1	I
Davis et al. (2003)	21	6.09–6.64	NoGo N2 Amp	GNG	Commission Errors	1	1
Epstein et al. (2014)	ب %	8-12	NoGo N2 Amp	GNG	Commission Errors		1
Felver (2013)	47	9-12 13 16	NoGo N2 Amp	5NG	1	ANT Conflict Task	
Gow et al. (2012) Grahell et al (2017)	011 %	1 2-10 4 27-5 67	NoGo N2 Amp	CNC CNC	1 1	1 1	ADHD Fyt Symptoms
Hum et al., 2013	89	8-12	NoGo N2 Amp	GNG	1	1	Int Symptoms
Inoue et al. (2010)	24 8	9-13	NoGo N2 Amp	GNG	1	I	ADHD
Johnstone et al., 2007a,	24	7.5-11.92	NoGo N2 Amp	GNG	Percent Correct (NoGo)	1	1
2007b							
Klymkiw et al. (2017)	67	11-17	NoGo N2 Amp	GNG	1	1	ADHD; Comorbid ADHD and Anxiety
Lahat et al. (2009)	37	4.1–6	NoGo N2 Amp	GNG	Accuracy Cost	1	1
Lamm et al. (2014)	106	7.28-8.47	NoGo N2 Amp	GNG	Percent Correct (NoGo)		Social Reticence; Behavioral Inhibition
Lamm et al. (2006)	ŝ	7.17–16.75	NoGo N2 Amp	GNG	1	Iowa Gambling Task, Stroop Task, Delay Discounting Backwards Digit Snan	I
Lónez-Martín et al. (2015)	48	8-13	NoGo N2 Amp	GNG	1	Discomiting, Dackwarus Digit Opan	ADHD
Petersen (2015)	87	2.5-3.5	NoGo N2 Amp	GNG	Percent Correct (NoGo)	Bear/Dragon, Shape Stroop, Grass-Snow	Ext Symptoms
Ruberry et al. (2017)	72	4.5-5.58	Go-NoGo N2 Dif	GNG	Percent Correct (NoGo)	NEPSY, Bear/Dragon, Day-Night, DCCS	•
			Waveform				
Shephard (2013)	43	9–17	NoGo N2 Amp	GNG	1	1	ADHD; Comorbid Tourette's and ADHD
Stieben et al. (2007)	37	8–12	NoGo N2 Amp	GNG	1	1	Ext Symptoms; Comorbid Ext and Int
Trinkl et al. (2015)	5 5	11-18	NoGo N2 Amp	GNG	1 0	I	Experience-Seeking; Disinhibition
Iropper (2009)	80.5	8-12		CNC CNC	Commission Errors	1	1
Without et al. (2016)	2 V	8.U8-9.42 0 1 9	Noco N2 Amp	5ND	Percent Correct (NOGO)	Attentional Founding/ Attentional Chifting/	[nottontive Dobariana] / Urmanative Immulaive
(2009) (2009)	3		diny zu anal			Persistence/ Impulsivity	Behaviors
Wiersema et al. (2006)	37	7–13	NoGo N2 Amp	GNG	1		ADHD
Woltering et al. (2011)	95	8-12	NoGo N2 Amp	GNG	1	1	Ext Symptoms
Zhang et al. (2016)	18	11.65-13.07	NoGo N2 Amp	GNG	1	1	Depressive Symptoms
Albrecht et al. (2005)	40	8-14	Stop-Go Dif Waveform/	SST	SSRT	1	ADHD; ODD or CD; comorbid ADHD and either
Rerner et al (2012)	18	5 5 00	Stop N2 Amp Stop N2 Amp	CCT			ODD of CD ADHD Symmetry
Dimoska et al. (2013)	26	7 42-12 92	Ston N2 Amn	LSS	- Failed Ston Trials vs Successful Ston Trials		
Janssen et al. (2015)	6	7-14	Stop N2 Amp	SST	Failed Stop Trials vs Successful Stop Trials	. 1	ADHD
Janssen et al. (2016)	81	7–13	Stop N2 Amp	SST	SSRT	1	-
Johnstone et al., 2007a,	24	7.5-11.92	Stop N2 Amp	SST	SSRT; Percentage of Successful Inhibition	1	1
20076	;				Trials		
Johnstone et al., 2007a, 2007b	88	8–14	Stop N2 Amp	SST	I	1	ADHD (Inattentive Subtype)
Liotti et al. (2007)	99	9–15	Stop N2 Amp	SST	Failed Stop Trials vs Successful Stop Trials	1	ADHD (Combined Subtype)
Lo et al. (2013)	20	5.1-6.9	Stop N2 Amp	SST	Failed Stop Trials vs Successful Stop Trials	1	1
Pliszka et al. (2000)	20	10.4–12.2			Inhibition Slope; Percentage of Successful Inhibition Trials	I	ADHD

(continued on next page)

Study	Z	Age Range (years)*	Predictor	Task	Outcomes		
					Neuropsychological Inhibition	Trait Self-Regulation	Psychopathology
Senderecka et al. (2012)	40	6.9–12.3	Stop N2 Amp	SST	Failed Stop Trials vs Successful Stop Trials; SSRT	1	ADHD (Combined Subtype)
Shen et al. (2011)	28	6.42 - 10.33	Stop N2 Amp	SST	Failed Stop Trials vs Successful Stop Trials	1	ADHD
Banaschewski et al. (2004)	64	8-14	NoGo N2 Amp	CPT	I	1	ADHD; ODD or CD; comorbid ADHD and either
							ODD or CD
Baving et al. (2004)	38	11-11.9	NoGo N2 Amp	CPT	Commission Errors	1	Anxiety
Fallgatter et al. (2004)	35	7.17-11.8	NoGo N2 Amp	CPT	1	1	ADHD
Jonkman et al. (2003)	12	9–10	NoGo N2 Amp	CPT	Percent Impulsive Errors/Percent False	I	1
					Alarms (NoGo/Total)		
Lawrence et al. (2005)	36	8.25-13.33	NoGo N2 Amp	CPT	1	1	ADHD
Overtoom et al. (1998)	26	6-14	NoGo N2 Amp	CPT	1	1	ADHD
Rommel et al. (2017)	204	11.90 - 25.90	NoGo N2 Amp	CPT	1	1	ADHD
Spronk et al. (2008)	27	5.90-7.70	NoGo N2 Amp	CPT	1	1	ADHD
Tye et al. (2014)	81	8-13	Go – NoGo N2 Dif	CPT	Omission Errors		ADHD/ Social Communication/ Inattentive
			Waveform				Symptoms/ Hyperactive Symptoms
			Waveloun				oymproms/ riyperacuve

Table 1 (continued)

reaction SSRT = stop signalTask, Signal = Stop Note: Amp = amplitude, GNG = Go/NoGo, ADHD = Attention Deficit Hyperactivity Disorder, Ext = externalizing, Int = internalizing, Dif = difference, SST For studies that did not report an age range, age range was computed as the mean age plus or minus the standard deviation of the age. ODD = Oppositional Defiant Disorder, CD = Conduct Disorder, CPT = Continuous Performance Task.

Neuroscience and Biobehavioral Reviews 103 (2019) 200-215

2.1.4. Outliers

As extreme effect size values are known to have a disproportionate influence on meta-analytic effect size combination, outliers were examined prior to combination (Hedges and Olkin, 1985). Outliers were defined using Tukey's method as those effect sizes outside of the interquartile range (Tukey, 1977). Using this method, no effect sizes were identified as outliers.

2.2. Effect size analysis

Across studies, associations between the N2 and the three behavioral phenotypes (behavioral response inhibition, trait self-regulation, and psychopathology) were typically presented in one of several forms: tests of homogeneity (i.e., t- or F- tests), correlation or regression tests, or Cohen's d. To determine if there was a significant association between the N2 and each phenotype of interest, the relevant information extracted from each subsample was converted into a common effect size metric. As examined phenotypic outcomes included scales on a variety of metrics, all effect sizes were converted into a standardized mean difference (SMD) value, in the form of Hedges' g. Hedges' g is an estimator of effect size that is appropriate for use when combining effect sizes derived from scales on different metrics (Hedges, 1982). However, Hedge's g is known to be positively biased when calculating effect sizes from small samples, and as such a correction factor was applied to reduce this positive bias (Durlak, 2009). Effect sizes were calculated using the "esc" package (Lüdecke, 2017) available for the statistical software program R (R Core Team, 2014).

For extracted effects, *SMD* values were calculated such that positive *g*-indexes indicated that larger (more negative) NoGo N2 amplitudes or larger difference waveform were associated with better task performance, better self-regulatory skills, or less psychopathology. Negative *g*-indexes indicated that larger (more negative) NoGo N2 amplitudes or larger difference waveforms were associated with worse task performance, poorer self-regulatory skills, or more psychopathology.

Once each effect size was represented as a SMD, they were combined into an overall mean effect size separately for each of the three behavioral phenotypes of interest using a fixed-effect model. Prior to combination, each effect size was weighted by its inverse variance, so that larger subsamples were assumed to more precisely estimate the population's effect size. Once each mean effect size was calculated, we explored the homogeneity of the effect size distribution. A homogeneous sample is one in which each effect size is an estimate of the same population effect size, and differences between effect sizes are attributable solely to subject-level sampling error. Homogeneity of effect sizes is an assumption of the fixed-effect model, so to examine if this assumption is tenable, we conducted a Q test of homogeneity, which quantifies the total amount of variance in effect size across cases (Cochran, 1954). In the Q test, if the null hypothesis is retained, the sample is homogeneous, and the fixed-effects model is appropriate. If the null hypothesis is rejected, there exists heterogeneity of effect sizes attributable to sources beyond subject-level sampling error (e.g., between-study differences), violating the assumptions of the fixed-effects model

To explore heterogeneity in our sample, we used a mixed-effects model, in which any heterogeneity present beyond sampling error was modeled to include both systematic components and random, unmeasured components. Our choice to use a mixed-effects model (instead of the more traditional random-effects model) was motivated by our desire to model heterogeneity as stemming from both systematic and random components and based on recommendations of Lipsey and Wilson (2001). The mixed-effects model allowed us to quantify heterogeneity stemming from systematic, between-study differences (e.g., age differences in the samples included in studies), while still including a random component that accounted for the residual variance remaining after the systematic differences were examined (Lipsey and Wilson, 2001). Estimates for both the fixed- and mixed-effects models

.

were generated using restricted maximum likelihood estimation (Viechtbauer, 2010).

In the mixed-effects model, systematic, between-study differences are explored as moderators. Plausible moderators examined in the current analysis included both categorical moderators: (1) task type (i.e., GNG, CPT, or SST), (2) psychopathology type (i.e., externalizing, internalizing, or comorbid), (3) the way in which psychopathology was measured (i.e., diagnoses vs. continuous measures of symptoms), (4) task difficulty (i.e., non-modified version of ERP task vs. a more complex version of the task), (5) ERP analysis technique (peak picking vs. mean amplitude vs. temporo-spatial PCA), and (6) N2 quantification technique used in analysis (NoGo N2 amplitude vs. difference waveform [NoGo N2 amplitude - Go N2 amplitude]), as well as continuous moderators: (7) sample size, (8) mean age of the sample, (9) percentage of females in the sample, and (10) percentage of inhibition trials included in the task (for the CPT, we calculated the percentage of NoGo trials [a cue stimulus followed by a non-target stimulus] among total presented cue trials; summary information about the percentage of inhibition and activation stimuli in studies included in our study are presented in Supplemental Appendix S1). All moderators were examined in separate analyses, with categorical moderators examined using an analog of an analysis of variance (ANOVA) approach and continuous moderators examined using an analog of a regression approach (Lipsey and Wilson, 2001; Viechtbauer, 2010).

Finally, the presence of bias in effect size estimates was investigated. Despite efforts to include as many studies as possible, there were a number of studies from which the required descriptive statistics could not be collected. Hence, it is possible that the "missing" cases, which included studies that did not report descriptive statistics as well as studies that were missing due to more traditional publication bias mechanisms, could be biasing the results of the effect size analysis. To examine potential bias due to this missingness, a funnel plot was generated and examined for asymmetry. Funnel plots graphically represent effect size by precision estimates for each case, with the most precise studies (likely the largest studies) located at the top of the funnel and the least precise studies (the smallest studies) located at the edges of the funnel. Funnel plot asymmetry is thought to indicate the presence of bias in a sample that is due to missing information, and is based on the assumption that each study's mean difference will be distributed symmetrically around the true standardized mean difference (Duval & Tweedy, 2000; Cooper, 2009; Viechtbauer, 2010). To determine if the funnel plot was asymmetric, a linear regression approach to measuring the asymmetry in the funnel plot was used (Egger et al., 1997; Viechtbauer, 2010). Significant results (in which the intercept of the regression line describing the association between effect size and precision [standard error] shows significant deviation from zero) suggest asymmetry, and indicate the need to adjust estimates based on the presence of bias. Non-significant results (in which the intercept of the regression line describing the association between effect size and precision [standard error] is not different from zero) indicate no need to adjust estimates. If asymmetry was detected, the trim and fill method would be used to adjust the overall effect size estimates (Duval and Tweedie, 2000; Cooper, 2009). As a another technique for examining publication bias in our sample, we also calculated a fail-safe N using the Orwin method, which quantifies the number of studies with null results that, if included in the effect size analysis, would reduce the observed average effect size to half of the current value (Orwin, 2019).

All meta-analytic models were fitted using the "metafor" (Viechtbauer, 2010) and "meta" (Schwarzer, 2007) packages available for the statistical software program R (R Core Team, 2014).

3. Results

3.1. Study characteristics

Across the 83 included subsamples, data from 3738 children (1118 female) were included in analyses. The average sample size, across

subsamples, was 45.03 children (SD = 37.95 children, range: 8–204). Effect size and any follow-up heterogeneity analyses are presented separately for each behavioral phenotype. Because so few subsamples (n = 6) could be included in the self-regulation analysis, we were underpowered to detect the association between the N2 and trait self-regulation (statistical power: .06). As such, we did not carry out this analysis.

3.2. Behavioral response inhibition

3.2.1. Effect size analysis

Across the 34 subsamples included in the behavioral response inhibition analysis, data from 1407 children (508 females) were included. The combined effect size from the fixed-effects model indicated an overall significant, negative effect size (SMD = -0.13, 95% CI: -0.23 to -0.03, p = .009). The combined, negative effect size indicates that a larger (more negative) NoGo N2 amplitude or difference waveform is associated with worse behavioral performance on tasks eliciting the N2. Of note, this finding is in the opposite direction of *a priori* expectations based on the literature. Fig. 2 includes a forest plot describing the various effect sizes that were included in the behavioral response inhibition analysis. Forest plots of this association, separated by task type, are presented in Supplemental Appendix S2.

The Q statistic resulting from the fixed-effect model surpassed the threshold for significance, suggesting that there was heterogeneity across effect sizes, more than what would be expected by chance or could be accounted for by sampling error alone (Q[33] = 95.44, p < .001). Given this substantial amount of heterogeneity, we used a mixed-effects modeling approach to explore several plausible moderators that might explain a portion of this heterogeneity. The moderator analyses are presented in Tables 2 and 3. Moderator analyses for categorical moderators, including task type (i.e., GNG, CPT, or SST), task difficulty (i.e., non-modified version of ERP task vs. a more complex version of the task), and ERP analysis technique (peak picking vs. mean amplitude vs. temporo-spatial PCA), are presented in Table 2. Moderator analyses for continuous moderators, including age of sample and percentage of inhibition trials included in the task, are presented in Table 3. Of the categorical moderators examined, the task type variable significantly moderated the association between the N2 and behavioral response inhibition. This finding suggests that the association between the N2 and behavioral response inhibition was larger and more negative when the N2 was elicited from the SST (compared to the GNG or CPT). Although task type accounted for 17.77% of the heterogeneity in effect size, a significant portion of heterogeneity remained unaccounted for (Q[31] = 76.25, p < .001). None of the other examined categorical or continuous moderators significantly moderated the association between the N2 and behavioral response inhibition.

Examination of sample bias in the behavioral response inhibition analysis indicated that the funnel plot of the sample was symmetrical (t [32] = 1.74, p = .09), suggesting no need to adjust this estimate for bias. The fail-safe N for the task analysis was calculated to be 34 studies.

3.3. Psychopathology

3.3.1. Effect size analysis

Across the 43 subsamples included in the overall psychopathology analysis, data from 2065 children (529 female) were included. The combined effect size from the fixed-effects model indicated an overall positive but non-significant effect size (*SMD* = 0.05, 95% CI: -0.04 to 0.14, p = .28). This finding suggests that there was no overall association between the N2 and psychopathology. Fig. 3 includes a forest plot describing the various effect sizes that were included in the overall psychopathology analysis. Forest plots of this association, separated by task, are presented in Supplemental Appendix S2.

The Q statistic resulting from the fixed-effect model surpassed the threshold for significance, suggesting the presence of a large amount of

between-study variance in effect size (Q[42] = 136.47, p < .0001). Given this substantial amount of heterogeneity, we used a mixed-effects modeling approach to explore several plausible moderators that might explain a portion of this heterogeneity. The moderator analyses are presented in Tables 2 and 3, with categorical moderators presented in Table 2, and continuous moderators presented in Table 3. Of primary importance to the current study, the psychopathology type variable (classifying studies as broadly assessing either externalizing or internalizing problems) significantly moderated the association between the N2 and psychopathology. Findings suggest that studies focusing on externalizing problems (including ADHD, conduct problems, and comorbid diagnoses that include externalizing problems) showed a larger, more positive association between the N2 and psychopathology (compared to studies focusing on internalizing problems). Studies focusing on internalizing problems showed a negative, but non-significant association. When a fixed-effects model was estimated separately for externalizing studies (i.e., a model that only included studies focusing on externalizing problems, including 38 subsamples and data from 1760 children [406 female]), the results suggest an overall positive, significant effect size (SMD = 0.13, 95% CI: -0.04 to 0.23, p = .01). A forest plot of only the studies focusing on externalizing problems is included in Fig. 4. Although the psychopathology type moderator accounted for 13.92% of the heterogeneity in effect size, a significant portion of heterogeneity in effect size remained unaccounted for after including this moderator (Q[41] = 117.88, p < .001). As a follow-up to these findings, we conducted two sub-analyses of the studies focusing on externalizing problems, examining the type of externalizing problem (i.e., ADHD vs. ODD/CD - excluding any cases which included subsamples with comorbid diagnoses) and the way in which psychopathology was measured (i.e., diagnoses vs. continuous measures of symptoms) as categorical moderators of the association between the N2 and psychopathology. These moderator analyses are presented in Table 2. Neither moderator significantly moderated the association between the N2 and psychopathology.

Of the other categorical moderators examined in the overall psychopathology analysis, two other variables moderated the association between the N2 and psychopathology: (1) task type and (2) N2 quantification technique. The moderation by task type was a trend-level finding, suggesting that the association between the N2 and psychopathology was larger when the N2 was elicited from a CPT (compared to the GNG or SST). However, as this moderator was only marginally significant, it will not be interpreted further. The moderation by N2 quantification technique suggested that the association between the N2 and overall psychopathology was larger when N2 amplitudes were used in analysis (a trend level association), and the association between the N2 and overall psychopathology was smaller (in the opposite direction of *a priori* expectations) when a difference waveform was used in analysis. However, given the marginally significant finding with N2 amplitudes, and the fact that only 3 samples, all from the same study, included in analysis used a difference waveform making us underpowered to examine this moderator effect, these findings were considered exploratory, and will not be interpreted further. None of the other examined categorical or continuous moderators significantly moderated the association between the N2 and overall psychopathology.

Examination of sample bias in the overall psychopathology analysis indicated that the funnel plot of the sample was symmetrical (t [41] = 1.23, p = .22), suggesting no need to adjust this estimate for bias. The fail-safe N for the psychopathology analysis was calculated to be 43 studies.

3.4. Statistical power

We investigated the statistical power that our meta-analysis had to detect the focal effects of interest. For the behavioral response inhibition analysis, our statistical power was .68, suggesting that we were adequately powered to detect the meta-analytic association between the N2 and behavioral response inhibition. Our statistical power for the externalizing problems analysis, was .79, suggesting that we were also adequately powered to detect the association between the N2 and externalizing problems.

4. Discussion

The current meta-analysis examined the association between the N2 ERP component and a variety of behavioral phenotypes that the N2 has been theorized to index. The child N2 literature is characterized by a pattern of inconsistent and sometimes contradictory findings. As such, the literature is in need of research that clarifies the nature of the childhood N2 and its external correlates to inform its functional significance. To fill this gap, the current study meta-analyzed this literature in order to examine the association of the N2 component with three



Fig. 2. Forest plot of studies included in behavioral response inhibition analysis.

Table 2

Moderation a	analysis of	categorical	moderators	across	behavioral	phenotypes.

Categorical Moderators	k	Estimate (SMD)	SE	p-value
Behavioral Response Inhibition				
Task type				
GNG Task	21	0.08	0.12	.51
SST	10	-0.39	0.20	.05
CPT	3	-0.13	0.33	.69
Task Difficulty				
Standard	24	-0.06	0.11	.63
Complex	10	-0.002	0.22	.99
ERP Analysis Method				
Peak-picking	17	-0.10	0.14	.48
Mean Amplitude	13	-0.04	0.21	.86
N2 Quantification Technique				
NoGo N2 Amplitude	32	-0.08	0.10	. 45
Difference Waveform	2	0.25	0.37	.50
Psychopathology				
Task type				
GNG Task	21	-0.03	0.12	.78
SST	12	0.19	0.22	.37
CPT	9	0.41	0.23	.08
Psychopathology Type				
Internalizing	5	-0.36	0.24	.15
Externalizing	38	0.53	0.26	.04
Psychopathology Measurement				
Diagnostic categories	29	0.05	0.11	.41
Symptom scores	13	0.09	0.21	.81
Task Difficulty				
Standard	34	0.11	0.10	.28
Complex	8	-0.05	0.24	.84
ERP Analysis Method				
Peak-picking	21	0.15	0.14	.27
Mean Amplitude	17	-0.04	0.21	.84
N2 Quantification Technique				
NoGo N2 Amplitude	40	0.17	0.09	.06
Difference Waveform	3	-1.24	0.38	0.001
Externalizing Studies Only				
Externalizing Type				
ADHD	23	0.22	0.13	.10
ODD/CD	4	-0.03	0.37	.93
Psychopathology Measurement				
Diagnostic categories	28	0.11	0.12	.35
Symptom scores	10	0.25	0.23	.29

Note: SMD = Standardized Mean Difference, GNG = Go/NoGo, SST = Stop Signal Task, CPT = Continuous.

Performance Task, ERP = Event-Related Potential, ADHD = Attention Deficit Hyperactivity Disorder.

ODD = Oppositional Defiant Disorder, CD = Conduct Disorder.

In the examination of categorical moderators, for the purpose of this analysis, the first listed category was examined as the baseline ("reference") *SMD*, and the subsequent estimate values reflect how much larger the *SMD* for the next listed moderator category was when compared to the baseline. For example, for the task type moderator of the neuropsychological inhibition analysis, the *SMD* for the GNG task studies was 0.08, while the *SMD* for the SST was -0.31 (0.08 - 0.39), and the *SMD* for the CPT was -0.05 (0.08 - 0.13).

Table 3					
Moderation analysis	of continuous	moderators	across	behavioral	phenotypes

Continuous Moderators	k	β	SE	p-value	R ²
Behavioral Response Inhibitio	n				
Sample Size	34	-0.001	0.002	.74	0.00
Mean Age	34	0.02	0.04	.69	0.00
Percent Female	30	-0.003	0.004	.43	0.00
NoGo Percentage	33	0.01	0.01	.12	0.09
Psychopathology					
Sample Size	43	-0.001	0.002	.61	0.00
Mean Age	43	-0.02	0.03	.60	0.00
Percent Female	43	-0.002	0.004	.67	0.00
NoGo Percentage	43	-0.008	0.01	.34	0.00

behavioral phenotypes: behavioral response inhibition, trait self-regulation, and externalizing versus internalizing psychopathology.

4.1. The N2 and behavioral response inhibition

The current study found that a larger (more negative) NoGo N2 amplitude or difference waveform was associated with worse behavioral response inhibition as quantified during the task used to elicit the N2. Notably, this finding is in the opposite direction of a priori expectations based on the adult literature (Falkenstein et al., 1999; Fogarty et al., 2018; Woltering et al., 2013). Given that our metaanalysis was adequately powered to detect the association between the N2 and behavioral response inhibition, we have increased confidence in this seemingly contradictory finding. There are several possible explanations for these findings. First, as described earlier, researchers have suggested that the N2 is an index of conflict monitoring. When conflicts are detected, this may trigger the use of alternative, compensatory attentional processes (Botvinick, 2007), which could, in turn, lead to increases in observed N2 amplitudes. As such, a larger N2 amplitude may reflect compensatory processing of relevant stimulus information, and might be associated with worsened task performance, because children with the greatest need for compensatory processing may struggle the most with completing the task. Alternatively, this seemingly contradictory finding could be due to the particularities of the N2 component elicited from children, namely the tendency for the amplitude of the N2 to decrease across childhood (Hoyniak, 2017). Our findings of an overall negative association between the N2 and behavioral response inhibition may reflect the tendency for N2 amplitudes to decrease and inhibitory performance to increase across childhood. These findings are consistent with evidence of neural activity becoming more efficient and focalized on task-relevant activity across development (Durston et al., 2006), likely leading to smaller amplitudes as neural activity becomes focalized.

Although we examined mean age of the sample as a moderator and the effect was not significant, there are a number of reasons that this moderator may not have been an ideal examination of age. First, many samples included in the meta-analysis accrued subjects across large age ranges. As such, the mean age of the sample may be less meaningful than the age of individual participants for the purposes of analysis if we had been able to obtain raw participant-level data. If the N2 has a different association with behavioral response inhibition at different stages of development, examining this association across a wide range of ages (as was included in the current meta-analysis) could lead to imprecise conclusions about the nature of the association between the N2 and behavioral response inhibition. Additional research is needed to clarify the role of child age in the association between the N2 and behavioral response inhibition, and it is apparent that these studies must span across small developmental windows in order to improve our capacity to detect age-related effects on the N2. Corresponding with this possible developmental interpretation of our findings, it is also possible that children with worse behavioral response inhibition skills show larger, more negative N2 amplitudes, because they are compensating for functional impairments in regions associated with response inhibition by showing overactivation in task-related and task-unrelated regions. This would be consistent with the neural compensation hypothesis as applied to both aging adults and individuals with ADHD individuals' known processing deficits across certain neural circuits may show a pattern of over-action across relevant and alternative neural circuits (Fassbender and Schweitzer, 2006; Reuter-Lorenz and Cappell, 2008). However, given the limitations of spatial information provided by ERPs, especially in ERPs elicited from children, this hypothesis is difficult to examine.

The heterogeneity across effect sizes included in the behavioral response inhibition phenotype was significantly moderated by task type, with the SST most likely to generate inhibitory N2s associated with behavioral response inhibition. While all three tasks, the GNG task, the



Fig. 3. Forest plot of studies included in overall psychopathology analysis.



Fig. 4. Forest plot of studies included in psychopathology analysis that focus on externalizing problems only.

CPT, and the SST, are thought to assess response inhibition and the overall effect across tasks was significant, the findings of the current study suggest that the SST might be best suited to assess specific inhibitory skills. Evidence that the SST may be a better measure of specific inhibitory skills compared to the GNG and CPT is consistent with evidence in adults that the GNG does not consistently elicit prepotent motor activity unless it is fast-paced and the inhibition trials are relatively rare (Wessel, 2018). Additionally, it is also possible that the N2 is more reflective of action cancellation capacities (better assessed by the SST) than action restraint capacities (better assessed by the GNG task and the CPT).

As only one significant moderator of the heterogeneity across effects sizes was identified, there remains a significant amount of unaccounted for variability in effect size. We examined many of the plausible moderators as explanations for the heterogeneity, but other untested moderators clearly explain a portion of this variance. One possible factor contributing to this high level of heterogeneity across studies is intra-individual variability in ERP responses. Research suggests that high levels of intra-individual variability in ERP latencies are present in clinical populations, including individuals with ADHD (Bluschke et al., 2017), as well as in developmental populations. This intra-individual variability in ERP latencies across trials is a significant problem, because ERP data are averaged across available trials, leading to the overlap of adjacent ERP components, as well as smearing of the morphology of components (Ouyang et al., 2011). This intra-individual variability, which may be especially significant in our meta-analysis that included data from both developmental and clinical populations, may affect amplitude values of measured ERPs, contributing to variability across samples and studies. One possible solution to explore and mitigate intra-individual variability in ERP latencies is to use singletrial analyses, combined with techniques to decompose the EEG signal (e.g., the residue iteration decomposition approach; Ouyang et al., 2011). Incorporating these analytic strategies to existing and newly collected datasets will improve our capacity to both identify components as well as better explain their functional meaning.

4.2. The N2 and trait self-regulation

As only six studies would have been able to be combined in the selfregulation effect size analysis there was insufficient power to examine the association between the overall meta-analytic association between the N2 and trait self-regulation. Future studies focused on the N2 component should examine whether trait self-regulatory skills are associated with the N2, to determine whether the N2 is associated with broader self-regulation skills (in addition to response inhibition, in particular).

4.3. The N2 and psychopathology

The current study suggests that the overall meta-analytic effect size for the psychopathology phenotype was not significantly different from zero. However, given the substantial number of studies included in this effect size analysis, we were able to consider several plausible categorical and continuous variables as moderators of this effect size. Of the moderators examined, two were found to significantly moderate the association between the N2 component and psychopathology: the type of psychopathology examined in the study and inhibition trial percentage of the task used to elicit the N2.

The moderation by psychopathology type indicated that studies focusing on individuals with some form of externalizing symptoms were more likely to show an association between the N2 and psychopathology, perhaps reflecting the commonly accepted notion that poor behavioral regulation is more central to externalizing problems then internalizing problems. This association between the N2 and externalizing psychopathology emerged in the expected direction, such that smaller N2 amplitudes were associated with higher levels of externalizing psychopathology. These studies either compared the N2 elicited from children meeting the diagnostic category thresholds for externalizing psychopathology (i.e., ADHD, ODD, CD, or a combination of externalizing and internalizing symptoms) with healthy controls, or examined specific externalizing symptoms (e.g., behavior problems, attention problems, etc.) in a group of children. Of note, there was no difference in effect size between studies that examined psychopathology using diagnostic categorization schemes and studies that focused on specific symptoms.

Although non-significant, the moderation results also suggested that pure internalizing symptoms (i.e., depression, anxiety, or a combination of depression and anxiety) might be associated with the N2 in the opposite direction, such that larger N2 amplitudes were associated with more internalizing problems. Despite that this effect was moderately sized (SMD = -0.36, p = .15), relatively few studies (n = 5) were included in this analysis, and this may have influenced why these results did not surpass the traditional threshold for significance. A full interpretation of these results is not warranted given the null findings, however, it is worth mentioning that this finding provides support for hypotheses about the different function of behavioral regulation in externalizing vs. internalizing problems. It has been hypothesized that externalizing problems represent a deficit in behavioral regulation, whereas internalizing problems may result from over-control of certain behaviors (e.g., Lewis and Stieben, 2004; Murray and Kochanska, 2002). Behavioral regulation has been proposed as a transdiagnostic marker of risk for psychopathology, in that too much or too little regulation can lead to maladaptive outcomes, especially in the context of an imbalance with other systems, such as behavioral approach (Jonas and Kochanska, 2018). Our results, which were adequately powered, provide support for externalizing problems being accompanied by the neural signature of a smaller N2 amplitude. Our results also hint at the possibility of the opposite association with internalizing problems (i.e., a larger N2 associated with more internalizing psychopathology). An increase in the number of studies examining the N2 in association with internalizing symptoms will further clarify this association.

A sub-analysis of the externalizing problems moderator, examining externalizing problem type (i.e., ADHD vs. ODD/CD) as a categorical moderator, demonstrated that externalizing problem type did not moderate the association between the N2 and psychopathology. Although this moderator did not surpass the threshold for significance, the association between the N2 and psychopathology was noticeably larger in studies focusing on ADHD (compared to ODD/CD), perhaps reflecting that deficits in response inhibition may be more core to ADHD, than to ODD/CD. This could also be due to the fact that many more studies focused on ADHD only (23 studies) than on disruptive behavior problems (ODD/CD only; 4 studies). Many of the studies that did include measures of disruptive behavior problems focused on general externalizing psychopathology (including both attention and disruptive behavior problems), and could not be included in this subanalysis, making firm conclusions about this moderator difficult. Future studies should consider examining these subtypes of externalizing problems separately. For now, though, our findings suggest that the association between the N2 and externalizing psychopathology did not differ based on the type of externalizing problems examined.

4.4. The N2 as a possible endophenotype

With the increasing focus on studying psychopathology from multiple levels of analysis, from cellular and genetic studies to brain networks and psychophysiology to behavior and symptom clusters, there has been increased interest in identifying endophenotypes of psychiatric diseases. Endophenotypes are heritable, "unobservable," intermediary traits that signify disease liability, and mediate the association between genotype and phenotypic expressions of psychopathology (Beauchaine, 2009; Gottesman and Gould, 2003). Different fields have had varying levels of success in identifying endophenotypes. However, most of this research has focused on adults, often because the discovery of brain-behavior relationships is even more complex when considered from a developmental framework (Casey et al., 2014; Lenzenweger, 2013). Given the intermediary role of endophenotypes in the association between genetic liability and phenotypic expressions of disease, establishing that an endophenotype is early-appearing and associated with risk for psychopathology is an important extension of the endophenotype construct. An endophenotype can be neurophysiological, biochemical, endocrinological, neuroanatomical, cognitive, or neuropsychological in nature (Gottesman and Gould, 2003). Consistent with the endophenotype hypothesis, neurophysiological phenotypes are thought to be better than cognitive phenotypes (e.g., response inhibition) in clarifying the substrates of psychopathology because neurophysiological phenotypes are closer to the underlying etiology (Jonas and Markon, 2014).

In addition to the P3 (Iacono and Malone, 2011), another neurophysiological phenotype that may be an endophenotype, or more precisely a measure of an endophenotype, for externalizing disorders is the N2. The results of the current study suggest that the N2 component satisfies several conditions of an endophenotype for externalizing psychopathology: (1) the association of the biomarker with the disease phenotype, (2) discriminant validity with respect to other phenotypes (namely internalizing psychopathology), and (3) appearance early on in life. While there are additional criteria that must be satisfied in order to confirm the N2 component as an endophenotype (e.g., establishing that the endophenotype is heritable and state-independent; Gottesman and Gould, 2003), this study provides an important first step. Examining the N2 may be a fruitful avenue for researchers interested in examining the neural underpinnings of regulation in childhood, especially because the N2 may be particularly useful as a neural marker of risk for

externalizing problems.

An alternative theoretical perspective, however, is that the N2 is a putative biomarker of behavioral response inhibition, and that behavioral response inhibition (rather than the N2) is an endophenotype for externalizing psychopathology. In this framework, the neural processes measured by the N2, as a neurophysiological manifestation of the cognitive process of behavioral response inhibition, serves the purpose of a biomarker, but is not an endophenotype. Or, perhaps both behavioral response inhibition and the neural processes measured by the N2 are an endophenotype across multiple levels of analysis. Future research will need to clarify these theoretical perspectives to understand the role of the N2 in externalizing psychopathology.

4.5. Limitations

The caveat with all meta-analyses is that the results are only as good as the studies it includes. As in any field, there is variability in the quality of the studies included in the current meta-analysis. Like many studies using neuroimaging techniques (Button et al., 2013), and especially studies focusing on children, the studies we examined tended to include relatively few participants (average N = 45.03, but often represent accrual across various groups of participants examined independently [e.g., children with ADHD and controls]) and to ignore statistical power (Larson and Carbine, 2017). The focus on examining individual differences may compound the seriousness of the modest sample size issue. Additionally, although it was of theoretical interest to examine the association between the N2 and trait self-regulation, because only a few studies included measures of trait self-regulation, we were unable to examine this association. Future research will also benefit from examining the role of the N2 in other inhibitory-type tasks like the Flanker task.

Additionally, some researchers have proposed that the N2 is more likely to be an index of response conflict instead of response inhibition. Our meta-analysis does not specifically address this disagreement in the field, but it will be important for researchers interested in using ERPs to study response inhibition in childhood to consider this interpretation of the N2 in their research moving forward. Very few of the studies we included in this meta-analysis considered this alternative, conflict-related meaning of the N2, and the field would be improved by considering this alternative interpretation in future research. This could include using creative experimental designs to separate response inhibition and response conflict skills to examine their association with the N2 component elicited from children in separate analyses. Additionally, research suggests that the inhibitory P3, a positivity typically coupled with the N2 in tasks assessing response inhibition, may be a better index of response inhibition than the N2 (Wessel and Aron, 2015). Future studies with children should include an explicit examination of the P3 component.

4.6. Conclusions

The current study sought to clarify our understanding of the meaning of the N2 component elicited in childhood across three widely used tasks, the GNG task, the CPT, and the SST. Our findings suggest that a smaller N2 is associated with improved behavioral response inhibition, in the form of better behavioral performance on the GNG task, the CPT, or the SST. Additionally, a smaller N2 was associated with more externalizing psychopathology. The findings of the current study suggest that further examination of the N2 as a potential endophenotype for externalizing disorders is warranted.

Funding

This material is based upon work supported by a National Science Foundation Graduate Research Fellowship [Hoyniak; grant number 1342962].

Declaration of Competing Interest

None.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.neubiorev.2019.06. 011.

References¹

- *Albrecht, B., Banaschewski, T., Brandeis, D., Heinrich, H., Rothenberger, A., 2005. Response inhibition deficits in externalizing child psychiatric disorders: an ERP-study with the Stop-task. Behav. Brain Funct. 1, 1–14. https://doi.org/10.1186/1744-9081-1-22.
- Aron, A.R., 2007. The neural basis of inhibition in cognitive control. Neuroscientist 13, 214–228. https://doi.org/10.1177/1073858407299288.
- *Banaschewski, T., Brandeis, D., Heinrich, H., Albrecht, B., Brunner, E., Rothenberger, A., 2004. Questioning inhibitory control as the specific deficit of ADHD–evidence from brain electrical activity. J. Neural Transm. 111, 841–864. https://doi.org/10.1007/ s00702-003-0040-8.
- *Barac, R., Moreno, S., Bialystok, E., 2016. Behavioral and electrophysiological differences in executive control between monolingual and bilingual children. Child Dev. 87, 1277–1290. https://doi.org/10.1111/cdev.12538.
- *Barry, R.J., De Blasio, F.M., 2015. Performance and ERP components in the equiprobable go/no-go task: Inhibition in children. Psychophysiology 52, 1228–1237. https://doi. org/10.1111/psyp.12447.
- Bates, J.E., Schermerhorn, A.C., Petersen, I.T., 2014. Temperament concepts in developmental psychopathology. In: Rudolph, K., Lewis, M. (Eds.), Handbook of Developmental Psychopathology, 3rd ed. Springer, New York, pp. 311–329.
- *Baving, L., Rellum, T., Laucht, M., Schmidt, M., 2004. Attentional enhancement to NoGo stimuli in anxious children. J. Neural Transm. 111, 985–999. https://doi.org/10. 1007/s00702-004-0116-0.
- Beauchaine, T.P., 2009. Role of biomarkers and endophenotypes in prevention and treatment of psychopathological disorders. Biomark. Med. 3, 1–3. https://doi.org/10. 2217/17520363.3.1.1.
- Bechara, A., Damasio, A.R., Damasio, H., Anderson, S.W., 1994. Insensitivity to future consequences following damage to human prefrontal cortex. Cognition 50, 7–15. https://doi.org/10.1016/0010-0277(94)90018-3.
- Bekker, E.M., Kenemans, J.L., Verbaten, M.N., 2005. Source analysis of the N2 in a cued Go/NoGo task. Cogn. Brain Res. 22, 221–231. https://doi.org/10.1016/j.cogbrainres. 2004.08.011.
- *Benikos, N., Johnstone, S.J., 2009. Arousal-state modulation in children with AD/HD. Clin. Neurophysiol. 120, 30–40. https://doi.org/10.1016/j.clinph.2008.09.026.
- Berger, A., 2011. Self-regulation: Brain, Cognition, and Development. American Psychological Association, Washington, DC.
- *Berger, A., Alyagon, U., Hadaya, H., Atzaba-Poria, N., Auerbach, J.G., 2013. Response inhibition in preschoolers at familial risk for attention deficit hyperactivity disorder: a behavioral and electrophysiological stop-signal study. Child Dev. 84, 1616–1632. https://doi.org/10.1111/cdev.12072.
- *Bluschke, A., Roessner, V., Beste, C., 2016. Specific cognitive-neurophysiological processes predict impulsivity in the childhood attention-deficit/hyperactivity disorder combined subtype. Psychol. Med. 46, 1277–1287. https://doi.org/10.1017/ S0033291715002822.
- *Bluschke, A., Chmielewski, W.X., Mückschel, M., Roessner, V., Beste, C., 2017. Neuronal intra-individual variability masks response selection differences between ADHD subtypes—a need to change perspectives. Front. Hum. Neurosci. 11https://doi.org/ 10.3389/fnhum.2017.00329. 197-10.
- Bokura, H., Yamaguchi, S., Kobayashi, S., 2001. Electrophysiological correlates for response inhibition in a Go/NoGo task. Clin. Neurophysiol. 112, 2224–2232. https:// doi.org/10.1016/S1388-2457(01)00691-5.
- Botvinick, M.M., 2007. Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. Cogn. Affect. Behav. Neurosci. 7, 356–366. https://doi.org/10.3758/CABN.7.4.356.
- *Boucher, O., Jacobson, J.L., Burden, M.J., Dewailly, É., Jacobson, S.W., Muckle, G., 2014. Prenatal tobacco exposure and response inhibition in school-aged children: an event-related potential study. Neurotoxicol. Teratol. 44, 81–88. https://doi.org/10. 1016/j.ntt.2014.06.003.
- Braver, T.S., 2012. The variable nature of cognitive control: a dual mechanisms framework. Trends Cogn. Sci. 16, 106–113. https://doi.org/10.1016/j.tics.2011.12.010.
- Brooker, R.J., Bates, J.E., Buss, K.A., Canen, M.J., Dennis-Tiwary, T.A., Gatzke-Kopp, L.M., Klein, D.N., Kujawa, A., Hoyniak, C.P., Lahat, A., Lamm, C., Moser, J.S., Petersen, I.T., Tang, A., Woltering, S., Schmidt, L.A., 2019. Conducting Event-Related Potential (ERP) Research with Young Children: A Review of Components, Special Considerations, and Recommendations for Research on Cognition and Emotion (under review).
- Bunford, N., Evans, S.W., Wymbs, F., 2015. ADHD and emotion dysregulation among children and adolescents. Clin. Child Fam. Psychol. Rev. 18, 185–217. https://doi.

¹ Studies included in the meta-analysis are indicated with an *.

org/10.1007/s10567-015-0187-5.

- Burle, B., Vidal, F., Bonnet, M., 2004. Electroencephalographic nogo potentials in a nomovement context: the case of motor imagery in humans. Neurosci. Lett. 360, 77–80. https://doi.org/10.1016/j.neulet.2004.02.034.
- Button, K.S., Ioannidis, J.P., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S., Munafò, M.R., 2013. Power failure: why small sample size undermines the reliability of neuroscience. Nat. Rev. Neurosci. 14, 365–376. https://doi.org/10.1038/nrn3475.
- Carlson, S.M., Moses, L.J., 2001. Individual differences in inhibitory control and children's theory of mind. Child Dev. 72, 1032–1053. https://doi.org/10.1111/1467-8624.00333.
- Casey, B.J., Oliveri, M.E., Insel, T., 2014. A neurodevelopmental perspective on the research domain criteria (RDoC) framework. Biol. Psychiatry 76, 350–353. https://doi. org/10.1016/j.biopsych.2014.01.006.
- Chatham, C.H., Frank, M.J., Munakata, Y., 2009. Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. Proc. Natl. Acad. Sci. 106, 5529–5533. https://doi.org/10.1073/pnas.0810002106.
- *Chevalier, N., Kelsey, K.M., Wiebe, S.A., Espy, K.A., 2014. The temporal dynamic of response inhibition in early childhood: an ERP study of partial and successful inhibition. Dev. Neuropsychol. 39, 585–599. https://doi.org/10.1080/87565641.2014. 973497.
- Cochran, W.G., 1954. The combination of estimates from different experiments. Biometrics 10, 101–129. https://doi.org/10.2307/3001666.
- Cooper, H., 2009. Research Synthesis and Meta-Analysis: A Step-by-Step Approach. Sage Publications, Los Angeles, CA.
- Cooper, H., Hedges, L.V., Valentine, J.C., 2009. The Handbook of Research Synthesis and Meta-Analysis. Russell Sage Foundation, New York, NY.
- *Cragg, L., Fox, A., Nation, K., Reid, C., Anderson, M., 2009. Neural correlates of successful and partial inhibitions in children: an ERP study. Dev. Psychobiol. 51, 533–543. https://doi.org/10.1002/dev.20391.
- *Davis, E.P., Bruce, J., Snyder, K., Nelson, C.A., 2003. The X-trials: neural correlates of an inhibitory control task in children and adults. J. Cogn. Neurosci. 15, 432–443. https://doi.org/10.1162/089892903321593144.
- DeBoer, T., Scott, L.S., Nelson, C.A., 2007. Methods for acquiring and analyzing infant event-related potentials. In: De Haan, M. (Ed.), Infant EEG and Event-Related Potentials. Psychology Press, New York, NY, pp. 5–37.
- Dickersin, K., 2005. Publication bias: recognizing the problem, understanding its origins and scope, and preventing harm. In: Rothstein, H.R., Sutton, A.J., Borenstein, M. (Eds.), Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments. John Wiley & Sons, pp. 11–33.
- *Dimoska, A., Johnstone, S.J., Barry, R.J., Clarke, A.R., 2003. Inhibitory motor control in children with attention-deficit/hyperactivity disorder: event-related potentials in the stop-signal paradigm. Biol. Psychiatry 54, 1345–1354. https://doi.org/10.1016/ S0006-3223(03)00703-0.
- Durlak, J.A., 2009. How to select, calculate, and interpret effect sizes. J. Pediatr. Psychol. 34, 917–928. https://doi.org/10.1093/jpepsy/jsp004.
- Durston, S., Davidson, M.C., Tottenham, N., Galvan, A., Spicer, J., Fossella, J.A., Casey, B.J., 2006. A shift from diffuse to focal cortical activity with development. Dev. Sci. 9, 1–20. https://doi.org/10.1111/j.1467-7687.2005.00454.x.
- Duval, S., Tweedie, R., 2000. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 56, 455–463. https:// doi.org/10.1111/j.0006-341X.2000.00455.x.
- Eagle, D.M., Bari, A., Robbins, T.W., 2008. The neuropsychopharmacology of action inhibition: cross-species translation of the stop-signal and go/no-go tasks. Psychopharmacology 199, 439–456. https://doi.org/10.1007/s00213-008-1127-6.
- Egger, M., Smith, G.D., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. BMJ Br. Med. J. 315, 629–634. https://doi.org/10.1136/ bmj.315.7109.629.
- *Epstein, B., Shafer, V.L., Melara, R.D., Schwartz, R.G., 2014. Can children with SLI detect cognitive conflict? Behavioral and electrophysiological evidence. J. Speech Lang. Hear, Res. 57, 1453–1467. https://doi.org/10.1044/2014 JSLHR-L-13-0234.
- Hear. Res. 57, 1453–1467. https://doi.org/10.1044/2014_JSLHR-L-13-0234.
 Falkenstein, M., Hoormann, J., Hohnsbein, J., 1999. ERP components in Go/Nogo tasks and their relation to inhibition. Acta Psychol. (Amst) 101, 267–291. https://doi.org/10.1016/S0001-6918(99)00008-6.
- *Fallgatter, A.J., Ehlis, A.-C., Seifert, J., Strik, W.K., Scheuerpflug, P., Zillessen, K.E., et al., 2004. Altered response control and anterior cingulate function in attentiondeficit/hyperactivity disorder boys. Clin. Neurophysiol. 115, 973–981. https://doi. org/10.1016/j.clinph.2003.11.036.
- Fan, J., McCandliss, B.D., Sommer, T., Raz, A., Posner, M.I., 2002. Testing the efficiency and independence of attentional networks. J. Cogn. Neurosci. 14, 340–347. https:// doi.org/10.1162/089892902317361886.
- Fassbender, C., Schweitzer, J.B., 2006. Is there evidence for neural compensation in attention deficit hyperactivity disorder? A review of the functional neuroimaging literature. Clin. Psychol. Rev. 26 (4), 445–465. https://doi.org/10.1016/j.cpr.2006.01. 003.
- *Felver, J., 2013. Understanding Children's Self-Regulation: An Analysis of Measurement and Change in the Context of a Mindfulness-based Intervention. (Doctor of Philosophy). University of Oregon, Eugene, OR.
- Fogarty, J.S., Barry, R.J., De Blasio, F.M., Steiner, G.Z., 2018. ERP components and behavior in the auditory equiprobable go/no-go task: Inhibition in young adults. Psychophysiology.
- Folstein, J.R., Van Petten, C., 2008. Influence of cognitive control and mismatch on the N2 component of the ERP: a review. Psychophysiology 45, 152–170. https://doi.org/ 10.1111/j.1469-8986.2007.00602.x.
- Frye, D., Zelazo, P.D., Palfai, T., 1995. Theory of mind and rule-based reasoning. Cogn. Dev. 10, 483–527. https://doi.org/10.1016/0885-2014(95)90024-1.
- Gerstadt, C.L., Hong, Y.J., Diamond, A., 1994. The relationship between cognition and

action: performance of children 312–7 years old on a stroop- like day-night test. Cognition 53, 129–153. https://doi.org/10.1016/0010-0277(94)90068-X.

- Gonthier, C., Macnamara, B.N., Chow, M., Conway, A.R.A., Braver, T.S., 2016. Inducing proactive control shifts in the AX-CPT. Front. Psychol. 7. https://doi.org/10.3389/ fpsyg.2016.01822.
- Gottsman, I.I., Gould, T.D., 2003. The endophenotype concept in psychiatry: etymology and strategic intentions. Am. J. Psychiatry 160, 636–645. https://doi.org/10.1176/ appi.ajp.160.4.636.
- *Gow, R.V., Rubia, K., Taylor, E., Vallée-Tourangeau, F., Matsudaira, T., Ibrahimovic, A., Sumich, A., 2012. Abnormal centroparietal ERP response in predominantly medication-naive adolescent boys with ADHD during both response inhibition and execution. J. Clin. Neurophysiol. 29, 181–189. https://doi.org/10.1097/WNP. 0b013e31824e1025.
- *Grabell, A.S., Olson, S.L., Tardif, T., Thompson, M.C., Gehring, W.J., 2017. Comparing self-regulation-associated event related potentials in preschool children with and without high levels of disruptive behavior. J. Abnorm. Child Psychol. 45, 1119–1132. https://doi.org/10.1007/s10802-016-0228-7.
- Groom, M.J., Cragg, L., 2015. Differential modulation of the N2 and P3 event-related potentials by response conflict and inhibition. Brain Cogn. 97, 1–9. https://doi.org/ 10.1016/j.bandc.2015.04.004.
- Hedges, L.V., 1982. Estimation of effect size from a series of independent experiments. Psychol. Bull. 92, 490–499. https://doi.org/10.1037/0033-2909.92.2.490.
- Hedges, L.V., Olkin, I., 1985. Statistical Method for Meta-Analysis. Academic Press, Orlando, FL.
- Hoyniak, C., 2017. Changes in the NoGo N2 Event-Related Potential component across childhood: a systematic review and meta-analysis. Dev. Neuropsychol. 42, 1–24. https://doi.org/10.1080/87565641.2016.1247162.
- Hoyniak, C.P., Petersen, I.T., Bates, J.E., Molfese, D.L., 2018. The neural correlates of temperamental inhibitory control in toddlers. Philos. Trans. R. Soc. B 373. https:// doi.org/10.1098/rstb.2017.0160.
- *Hum, K.M., Manassis, K., Lewis, M.D., 2013. Neural mechanisms of emotion regulation in childhood anxiety. J. Child Psychol. Psychiatry 54, 552–564 doi: 0.1111/j.1469-7610.2012.02609.x.
- Huster, R.J., Eichele, T., Enriquez-Geppert, S., Wollbrink, A., Kugel, H., Konrad, C., Pantev, C., 2011. Multimodal imaging of functional networks and event-related potentials in performance monitoring. NeuroImage 56, 1588–1597. https://doi.org/10. 1016/j.neuroimage.2011.03.039.
- Huster, R.J., Enriquez-Geppert, S., Lavallee, C.F., Falkenstein, M., Herrmann, C.S., 2013. Electroencephalography of response inhibition tasks: Functional networks and cognitive contributions. Int. J. Psychophysiol. 87, 217–233. https://doi.org/10.1016/j. iipsycho.2012.08.001.
- Iacono, W.G., Malone, S.M., 2011. Developmental endophenotypes: indexing genetic risk for substance abuse with the P300 brain event-related potential. Child Dev. Perspect. 5, 239–247. https://doi.org/10.1111/j.1750-8606.2011.00205.x.
- *Inoue, Y., Inagaki, M., Gunji, A., Furushima, W., Okada, H., Sasaki, H., et al., 2010. Altered effect of preceding response execution on inhibitory processing in children with AD/HD: an ERP study. Int. J. Psychophysiol. 77, 118–125. https://doi.org/10. 1016/j.ijpsycho.2010.05.002.
- Insel, T.R., 2014. Mental disorders in childhood: shifting the focus from behavioral symptoms to neurodevelopmental trajectories. JAMA 311, 1727–1728. https://doi. org/10.1001/jama.2014.1193.
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D.S., Quinn, K., et al., 2010. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Acad. Psychiatry 167, 748–751. https://doi.org/10.1176/appi. ajp.2010.09091379.
- Isbell, E., Calkins, S.D., Cole, V.T., Swingler, M.M., Leerkes, E.M., 2019. Longitudinal Associations between Conflict Monitoring and Emergent Academic Skills: An Event-Related Potentials Study (in press). https://doi.org/10.1002/dev.21809.
- *Janssen, T.W.P., Bink, M., Geladé, K., van Mourik, R., Maras, A., Oosterlaan, J., 2016. A randomized controlled trial investigating the effects of neurofeedback, methylphenidate, and physical activity on event-related potentials in children with Attention-Deficit/Hyperactivity disorder. J. Child Adolesc. Psychopharmacol. 26, 344–353. https://doi.org/10.1089/cap.2015.0144.
- *Janssen, T.W., Heslenfeld, D.J., van Mourik, R., Geladé, K., Maras, A., Oosterlaan, J., 2015. Alterations in the ventral attention network during the stop-signal task in children with ADHD: an event-related potential source imaging study. J. Atten. Disord. https://doi.org/10.1177/1087054715580847.
- *Johnstone, S.J., Barry, R.J., Clarke, A.R., 2007a. Behavioural and ERP indices of response inhibition during a Stop-signal task in children with two subtypes of Attention-Deficit Hyperactivity Disorder. Int. J. Psychophysiol. 66, 37–47. https:// doi.org/10.1016/j.ijpsycho.2007.05.011.
- Johnstone, S.J., Clarke, A.R., 2009. Dysfunctional response preparation and inhibition during a visual Go/Nogo task in children with two subtypes of attention-deficit hyperactivity disorder. Psychiatry Res. 166, 223–237. https://doi.org/10.1016/j. psychres.2008.03.005.
- *Johnstone, S.J., Dimoska, A., Smith, J.L., Barry, R.J., Pleffer, C.B., Chiswick, D., Clarke, A.R., 2007b. The development of stop-signal and Go/Nogo response inhibition in children aged 7–12 years: performance and event-related potential indices. Int. J. Psychophysiol. 63, 25–38. https://doi.org/10.1016/j.ijpsycho.2006.07.001.
- Jonas, K.G., Kochanska, G., 2018. An imbalance of approach and effortful control predicts externalizing problems: support for extending the dual-systems model into early childhood. J. Abnorm. Child Psychol. https://doi.org/10.1007/s10802-018-0400-3.
- Jonas, K.G., Markon, K.E., 2014. A meta-analytic evaluation of the endophenotype hypothesis: effects of measurement paradigm in the psychiatric genetics of impulsivity. J. Abnorm. Psychol. 123, 660–675. https://doi.org/10.1037/a0037094.
- *Jonkman, L., Lansbergen, M., Stauder, J., 2003. Developmental differences in behavioral

and event-related brain responses associated with response preparation and inhibition in a go/nogo task. Psychophysiology 40, 752–761. https://doi.org/10.1111/ 1469-8986.00075.

Jonkman, L., Sniedt, F., Kemner, C., 2007. Source localization of the Nogo-N2: a developmental study. Clin. Neurophysiol. 118, 1069–1077. https://doi.org/10.1016/j. clinph.2007.01.017.

*Klymkiw, D.F., Milligan, K., Lackner, C., Phillips, M., Schmidt, L.A., Segalowitz, S.J., 2017. Does anxiety enhance or hinder attentional and impulse control in youth with ADHD? An ERP analysis. J. Atten. Disord. https://doi.org/10.1177/ 1087054717707297.

Kujawa, A., Hajcak, G., Torpey, D., Kim, J., Klein, D.N., 2012. Electrocortical reactivity to emotional faces in young children and associations with maternal and paternal depression. J. Child Psychol. Psychiatry 53, 207–215. https://doi.org/10.1111/j.1469-7610.2011.02461.x.

*Lahat, A., Todd, R.M., Mahy, C.E.V., Lau, K., Zelazo, P.D., 2009. Neurophysiological correlates of executive function: a comparison of European-Canadian and Chinese-Canadian 5-year-old children. Front. Hum. Neurosci. 3, 72–100. https://doi.org/10. 3389/neuro.09.072.2009.

*Lamm, C., Walker, O.L., Degnan, K.A., Henderson, H.A., Pine, D.S., McDermott, J.M., Fox, N.A., 2014. Cognitive control moderates early childhood temperament in predicting social behavior in 7-year-old children: an ERP study. Dev. Sci. 17, 667–681. https://doi.org/10.1111/desc.12158.

*Lamm, C., Zelazo, P.D., Lewis, M.D., 2006. Neural correlates of cognitive control in childhood and adolescence: disentangling the contributions of age and executive function. Neuropsychologia 44, 2139–2148. https://doi.org/10.1016/j. neuropsychologia.2005.10.013.

Larson, M.J., Carbine, K.A., 2017. Sample size calculations in human electrophysiology (EEG and ERP) studies: a systematic review and recommendations for increased rigor. Int. J. Psychophysiol. 111, 33–41. https://doi.org/10.1016/j.ijpsycho.2016.06.015.

*Lawrence, C.A., Barry, R.J., Clarke, A.R., Johnstone, S.J., McCarthy, R., Selikowitz, M., Broyd, S.J., 2005. Methylphenidate effects in attention deficit/hyperactivity disorder: electrodermal and ERP measures during a continuous performance task. Psychopharmacology 183, 81–91. https://doi.org/10.1007/s00213-005-0144-y.

Lavric, A., Pizzagalli, D.A., Forstmeier, S., 2004. When 'go'and 'nogo'are equally frequent: ERP components and cortical tomography. Eur. J. Neurosci. 20, 2483–2488. https:// doi.org/10.1111/j.1460-9568.2004.03683.x.

Lenzenweger, M.F., 2013. Thinking clearly about the endophenotype–intermediate phenotype–biomarker distinctions in developmental psychopathology research. Dev. Psychopathol. 25, 1347–1357. https://doi.org/10.1017/S0954579413000655.

Lewis, M.D., Lamm, C., Segalowitz, S.J., Stieben, J., Zelazo, P.D., 2006. Neurophysiological correlates of emotion regulation in children and adolescents. J. Cogn. Neurosci. 18, 430–443. https://doi.org/10.1162/jocn.2006.18.3.430.

Lewis, M.D., Stieben, J., 2004. Emotion regulation in the brain: conceptual issues and directions for developmental research. Child Dev. 75, 371–376. https://doi.org/10. 1111/j.1467-8624.2004.00680.x.

Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P., et al., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann. Intern. Med. 151, W-65–W-94. https://doi.org/10.7326/0003-4819-151-4-200908180-00136.

Lipsey, M.W., Wilson, D.B., 2001. Practical Meta-Analysis. Sage Publications, Inc, Thousand Oaks, CA.

*Liotti, M., Pliszka, S.R., Perez, R., Luus, B., Glahn, D., Semrud-Clikeman, M., 2007. Electrophysiological correlates of response inhibition in children and adolescents with ADHD: influence of gender, age, and previous treatment history. Psychophysiology 44, 936–948. https://doi.org/10.1111/j.1469-8986.2007. 00568.x.

*Lo, Y.-H., Liang, W.-K., Lee, H.-W., Wang, C.-H., Tzeng, O.J.L., Hung, D.L., et al., 2013. The neural development of response inhibition in 5- and 6-year-old preschoolers: An ERP and EEG study. Dev. Neuropsychol. 38, 301–316. https://doi.org/10.1080/ 87565641.2013.801980.

*López-Martín, S., Albert, J., Fernández-Jaén, A., Carretié, L., 2015. Emotional response inhibition in children with attention-deficit/hyperactivity disorder: neural and behavioural data. Psychol. Med. 45, 2057–2071. https://doi.org/10.1017/ S0033291714003195.

Lüdecke, D., 2017. Esc: Effect Size Computation for Meta-analysis. R Package Version 0.4.0. URL:. https://CRAN.R-project.orge/package=esc.

Meyer, A., Hajcak, G., Torpey-Newman, D.C., Kujawa, A., Klein, D.N., 2015. Enhanced error-related brain activity in children predicts the onset of anxiety disorders between the ages of 6 and 9. J. Abnorm. Psychol. 124, 266–274. https://doi.org/10.1037/ abn0000044.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann. Intern. Med. 151, 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135.

Murray, K.T., Kochanska, G., 2002. Effortful control: factor structure and relation to externalizing and internalizing behaviors. J. Abnorm. Child Psychol. 30, 503–514. https://doi.org/10.1023/A:1019821031523.

Nieuwenhuis, S., Yeung, N., Cohen, J.D., 2004. Stimulus modality, perceptual overlap, and the go/no-go N2. Psychophysiology 41, 157–160. https://doi.org/10.1046/j. 1469-8986.2003.00128.x.

Nieuwenhuis, S., Yeung, N., Van Den Wildenberg, W., Ridderinkhof, K.R., 2003. Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. Cogn. Affect. Behav. Neurosci. 3, 17–26. https://doi.org/10.3758/CABN.3.1.17.

Orwin, R.G., 2019. A fail-safe N for effect size in meta-analysis. J. Educ. Stat. 8, 157–159. *Overtoom, C.C., Verbaten, M.N., Kemner, C., Kenemans, J.L., van Engeland, H., Buitelaar, J.K., et al., 1998. Associations between event-related potentials and measures of attention and inhibition in the Continuous Performance Task in children with ADHD and normal controls. J. Am. Acad. Child Adolesc. Psychiatry 37, 977–985. https://doi.org/10.1097/00004583-199809000-00018.

*Petersen, I.T., 2015. Neurophysiological Mechanisms in the Development of Externalizing Behavior Problems in Young Children. (Doctor of Philosophy). Indiana University, Bloomington, IN.

Petersen, I.T., Bates, J.E., Dodge, K.A., Lansford, J.E., Pettit, G.S., 2015. Describing and predicting developmental profiles of externalizing problems from childhood to adulthood. Dev. Psychopathol. 27, 791–818. https://doi.org/10.1017/ S0954579414000789.

*Pliszka, S.R., Liotti, M., Woldorff, M.G., 2000. Inhibitory control in children with attention-deficit/hyperactivity disorder: event-related potentials identify the processing component and timing of an impaired right-frontal response-inhibition mechanism. Biol. Psychiatry 48, 238–246. https://doi.org/10.1016/S0006-3223(00) 00890-8.

Polanczyk, G.V., Salum, G.A., Sugaya, L.S., Caye, A., Rohde, L.A., 2015. Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. J. Child Psychol. Psychiatry 56, 345–365. https://doi.org/10.1111/ jcpp.12381.

Proudfit, G.H., 2015. The reward positivity: from basic research on reward to a biomarker for depression. Psychophysiology 52, 449–459. https://doi.org/10.1111//psyp.. 12370.

R Core Team, 2014. R: A Language and Environment for Statistical Computing. URL. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Reed, M.A., Pien, D.L., Rothbart, M.K., 1984. Inhibitory self-control in preschool children. Merrill-Palmer Q. 30, 131–147.

Reuter-Lorenz, P.A., Cappell, K.A., 2008. Neurocognitive aging and the compensation hypothesis. Curr. Dir. Psychol. Sci. 17, 177–182. https://doi.org/10.1111/j.1467-8721.2008.00570.x.

- *Rommel, A.-S., James, S.-N., McLoughlin, G., Brandeis, D., Banaschewski, T., Asherson, P., Kuntsi, J., 2017. Association of preterm birth with Attention-Deficit/Hyperactivity Disorder–Like and wider-ranging neurophysiological impairments of attention and inhibition. J. Am. Acad. Child Adolesc. Psychiatry 56, 40–50. https://doi.org/10. 1016/j.jaac.2016.10.006.
- *Ruberry, E.J., Lengua, L.J., Crocker, L.H., Bruce, J., Upshaw, M.B., Sommerville, J.A., 2017. Income, neural executive processes, and preschool children's executive control. Dev. Psychopathol. 29, 143–154. https://doi.org/10.1017/S095457941600002X.

Schachar, R., Logan, G.D., Robaey, P., Chen, S., Ickowicz, A., Barr, C., 2007. Restraint and cancellation: multiple inhibition deficits in attention deficit hyperactivity disorder. J. Abnorm. Child Psychol. 35, 229–238. https://doi.org/10.1007/s10802-006-9075-2.

Schwarzer, G., 2007. Meta: an R package for meta-analysis. R News 7, 40–45. Schwarzer, G., Carpenter, J.R., Rücker, G., 2015. Meta-Analysis With R. Springer.

Schoemaker, K., Mulder, H., Deković, M., Matthys, W., 2013. Executive functions in preschool children with externalizing behavior problems: a meta-analysis. J. Abnorm. Child Psychol. 41, 457–471. https://doi.org/10.1007/s10802-012-9684-x.

*Senderecka, M., Grabowska, A., Szewczyk, J., Gerc, K., Chmylak, R., 2012. Response inhibition of children with ADHD in the stop-signal task: an event-related potential study. Int. J. Psychophysiol. 85, 93–105. https://doi.org/10.1016/j.ijpsycho.2011. 05.007.

*Shen, I.-H., Tsai, S.-Y., Duann, J.-R., 2011. Inhibition control and error processing in children with attention deficit/hyperactivity disorder: an event-related potentials study. Int. J. Psychophysiol. 81, 1–11. https://doi.org/10.1016/j.ijpsycho.2011.03. 015.

*Shephard, E., 2013. Electrophysiological Correlates of Learning and Cognitive Control in Children With Tics With and Without ADHD Symptoms (Doctor of Philosophy). University of Nottingham.

Simmonds, D.J., Pekar, J.J., Mostofsky, S.H., 2008. Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. Neuropsychologia 46, 224–232. https://doi.org/10.1016/j. neuropsychologia.2007.07.015.

Slaats-Willemse, D., Swaab-Barneveld, H., De Sonneville, L., Van Der Meulen, E., Buitelaar, J., 2003. Deficient response inhibition as a cognitive endophenotype of ADHD. J. Am. Acad. Child Adolesc. Psychiatry 42, 1242–1248. https://doi.org/10. 1097/01.chi.0000081825.25107.f2.

*Spronk, M., Jonkman, L., Kemner, C., 2008. Response inhibition and attention processing in 5-to 7-year-old children with and without symptoms of ADHD: an ERP study. Clin. Neurophysiol. 119, 2738–2752. https://doi.org/10.1016/j.clinph.2008.09.010.

Steele, V.R., Aharoni, E., Munro, G.E., Calhoun, V.D., Nyalakanti, P., Stevens, M.C., et al., 2013. A large scale (N= 102) functional neuroimaging study of response inhibition in a Go/NoGo task. Behav. Brain Res. 256, 529–536. https://doi.org/10.1016/j.bbr. 2013.06.001.

*Stieben, J., Lewis, M.D., Granic, I., Zelazo, P.D., Segalowitz, S., Pepler, D., 2007. Neurophysiological mechanisms of emotion regulation for subtypes of externalizing children. Dev. Psychopathol. 19, 455–480. https://doi.org/10.1017/ S0954579407070228.

Stroop, J.R., 1935. Studies of interference in serial verbal reactions. J. Exp. Psychol. 18, 643–662. https://doi.org/10.1037/h0054651.

Sutton, A.J., 2009. Publication Bias. The Handbook of Research Synthesis and Meta-Analysis. pp. 435–452.

Swick, D., Ashley, V., Turken, U., 2011. Are the neural correlates of stopping and not going identical? Quantitative meta-analysis of two response inhibition tasks. NeuroImage 56, 1655–1665. https://doi.org/10.1016/j.neuroimage.2011.02.070.

Thierry, G., 2005. The use of event-related potentials in the study of early cognitive development. Infant Child Dev. 14, 85–94. https://doi.org/10.1002/icd.353.

*Trinkl, M., Greimel, E., Bartling, J., Grünewald, B., Schulte-Körne, G., Grossheinrich, N.,

2015. Right-lateralization of N2-amplitudes in depressive adolescents: an emotional go/no-go study. J. Child Psychol. Psychiatry 56, 76–86. https://doi.org/10.1111/jcpp.12282.

- *Tropper, B., 2009. An Electrophysiological and Behavioral Examination of Cognitive Control in Children with Specific Language Impairment (Doctor of Philosophy). City University of New York.
- Tukey, J.W., 1977. Exploratory Data Analysis, 1 ed. Pearson, Reading, MA.
- *Tye, C., Asherson, P., Ashwood, K.L., Azadi, B., Bolton, P., McLoughlin, G., 2014. Attention and inhibition in children with ASD, ADHD and co-morbid ASD + ADHD: an event-related potential study. Psychol. Med. 44, 1101–1116. https://doi.org/10. 1017/S0033291713001049.
- Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. J. Stat. Softw. 36 (1-), 48.
- *Vuillier, L., Bryce, D., Szücs, D., Whitebread, D., 2016. The maturation of interference suppression and response inhibition: ERP analysis of a cued Go/Nogo task. PLoS One 11, e0165697. https://doi.org/10.1371/journal.pone.0165697.
- Wessel, J.R., 2018. Prepotent motor activity and inhibitory control demands in different variants of the go/no-go paradigm. Psychophysiology 55. https://doi.org/10.1111/ psyp.12871.
- Wessel, J.R., Aron, A.R., 2015. It's not too late: the onset of the frontocentral P3 indexes successful response inhibition in the stop-signal paradigm. Psychophysiology 52, 472–480. https://doi.org/10.1111/psyp.12374.

*Wiersema, J.R., Roeyers, H., 2009. ERP correlates of effortful control in children with

varying levels of ADHD symptoms: temperament and psychopathology. J. Abnorm. Child Psychol. 37, 327–336. https://doi.org/10.1007/s10802-008-9288-7.

- *Wiersema, R., Van Der Meere, J., Roeyers, H., Van, C., Baeyens, D., 2006. Event rate and event-related potentials in ADHD. J. Child Psychol. Psychiatry 47, 560–567. https:// doi.org/10.1111/j.1469-7610.2005.01592.x.
- *Woltering, S., Granic, I., Lamm, C., Lewis, M.D., 2011. Neural changes associated with treatment outcome in children with externalizing problems. Biol. Psychiatry 70, 873–879. https://doi.org/10.1016/j.biopsych.2011.05.029.
- Woltering, S., Liu, Z., Rokeach, A., Tannock, R., 2013. Neurophysiological differences in inhibitory control between adults with ADHD and their peers. Neuropsychologia 51, 1888–1895. https://doi.org/10.1016/j.neuropsychologia.2013.06.023.
- Young, S.E., Friedman, N.P., Miyake, A., Willcutt, E.G., Corley, R.P., Haberstick, B.C., Hewitt, J.K., 2009. Behavioral disinhibition: liability for externalizing spectrum disorders and its genetic and environmental relation to response inhibition across adolescence. J. Abnorm. Psychol. 118, 117–130. https://doi.org/10.1037/a0014657.
- *Zhang, W., Ding, Q., Chen, N., Wei, Q., Zhao, C., Zhang, P., et al., 2016. The development of automatic emotion regulation in an implicit emotional Go/NoGo paradigm and the association with depressive symptoms and anhedonia during adolescence. Neuroimage Clin. 11, 116–123. https://doi.org/10.1016/j.nicl.2016.01.018.
- Zhou, Q., Chen, S.H., Main, A., 2011. Commonalities and differences in the research on children's effortful control and executive function: a call for an integrated model of self-regulation. Child Dev. Perspect. 6, 112–121. https://doi.org/10.1111/j.1750-8606.2011.00176.x.

Supplementary Appendix S1

Of the GNG task and SST studies we examined, 8 used a 50(Go)/50(NoGo) split, 5 used a 66(Go)/33(NoGo) split, 13 used a 70(Go)/30(NoGo) split, 15 used a 75(Go)/25(NoGo) split, and 6 used a 80(Go)/20(NoGo) split. For the CPT task, at the suggestion of Reviewer 2, we calculated the proportion of Go to NoGo trials slightly differently, in that the total number of trials reflects just the total number of cued trials, divided into the percent that are cued-Go trials and the percent that are cued-NoGo trials. Coded in this way, 9 of the CPT studies had a 50(cued-Go)/50(cued-NoGo) split.



Author(s) and Year		SMD [95% CI]
Barac, Moreno, & Bialystok (2016) Barry & De Blasio (2015) Bluschke, Roessner, & Beste (2016) Boucher et al. (2014) Chevalier et al. (2014) Cragg et al. (2009) Davis et al. (2009) Davis et al. (2009) Davis et al. (2009) Laster et al. (2014) Epstein et al. (2014) Johnstone et al. (2007) Lahat et al. (2009) Lahat et al. (2009) Lamm et al. (2014) Petersen (2015) Ruberry et al. (2017) Tropper (2009) Tropper (2009) Vuillier et al. (2016)		0.00 [-0.50, 0.50] 0.77 [0.13, 1.41] 1.61 [0.85, 2.37] -0.20 [-0.49, 0.09] -0.10 [-0.82, 0.62] -1.26 [-2.11, -0.42] 0.24 [-0.47, 0.96] 0.00 [-0.86, 0.86] 0.00 [-0.86, 0.86] 0.00 [-0.84, 0.84] 0.00 [-0.98, 0.98] -1.12 [-1.99, -0.26] -0.26 [-1.14, 0.62] 0.22 [-0.73, 1.17] 0.00 [-0.38, 0.38] 0.56 [-0.12, 1.25] 0.04 [-0.50, 0.58] 0.47 [0.00, 0.94] -0.20 [-1.04, 0.64] 0.39 [-0.60, 1.38] -0.02 [-0.94, 0.90]
Wiersema & Roeyers (2009)	· · · · · · · · · · · · · · · · · · ·	0.16 [-0.61, 0.93]
Fixed-Effects Model	*	0.05 [-0.08, 0.19]
-2.5	-1 0 1	2.5
	SMD	

Forest plots for studies included in the Behavioral Response Inhibition analysis, separated by task type

Figure S1. Forest plot of Go/NoGo studies included in behavioral response inhibition analysis



Figure S2. Forest plot of Stop Signal task studies included in behavioral response inhibition analysis



Figure S3. Forest plot of Continuous Performance Task studies included in behavioral response inhibition analysis

Forest plots for studies included in the Psychopathology analysis, separated by task type



Figure S4. Forest plot of Go/NoGo studies included in the psychopathology analysis



Figure S5. Forest plot of Stop Signal task studies included in the psychopathology analysis



Figure S6. Forest plot of Continuous Performance task studies included in the psychopathology analysis