The Pennsylvania State University
The Graduate School
College of Health and Human Development

RELEVANCE OF THE P3B EVENT-RELATED POTENTIAL TO CHILDREN’S
SELF-REGULATION AND ACADEMIC PERFORMANCE:
FINDINGS FROM A DEVELOPMENTAL CASCADE MODEL

A Thesis in
Human Development and Family Studies

by
Cynthia J. Willner

© 2013 Cynthia J. Willner

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

August 2013
The thesis of Cynthia J. Willner was reviewed and approved* by the following:

Lisa M. Gatzke-Kopp  
Associate Professor of Human Development  
Thesis Advisor

Charles F. Geier  
Assistant Professor of Human Development

J. Douglas Coatsworth  
Professor of Human Development  
Professor-in-Charge, Graduate Program in Human Development and Family Studies

*Signatures are on file in the Graduate School.
ABSTRACT

Children’s ability to regulate their attention, emotions and behaviors at school entry is critical for setting the stage for positive school adjustment and optimal academic performance. Although our understanding of individual differences in the neural mechanisms underlying self-regulatory control is rapidly advancing, measures of neurophysiological functioning are rarely integrated into school-based research. One neurophysiological measure that has been shown to be relevant to self-regulation is the P3b event-related potential, a positive deflection in the EEG waveform that indexes selective attention allocation. Studies with older children and adolescents have shown that low P3b amplitudes are associated with greater impulsivity/distractibility, lower performance on IQ tests, poorer reading achievement, and increased risk for externalizing psychopathology. However, the developmental processes underlying these associations are not well understood. This paper uses a developmental cascade model to examine how individual differences in the amplitude of high-risk kindergarten (KG) children’s P3b event-related potentials, measured following target stimuli in a Go/No-Go task, predict their behavioral self-regulation and academic performance in KG, 1st, and 2nd grade. This study provides evidence that higher P3b amplitudes are associated with better behavioral self-regulation in the kindergarten classroom, which mediates a positive indirect effect of P3b amplitude on academic performance in 1st grade. Given that P3b amplitude indexes selective attention processes, these findings build on the scientific justification for interventions targeting young children’s attention skills in order to promote their self-regulation and academic achievement. P3b amplitude, or a closely-linked neurocognitive assessment, may also prove useful as a tool for identifying children to target for specific interventions and for assessing the potential impacts of interventions on the functioning of neural networks subserving selective attention.
# Table of Contents

List of Tables .............................................................................................................................................. v  
List of Figures .............................................................................................................................................. vi  
List of Abbreviations ................................................................................................................................. vii  
Acknowledgements ...................................................................................................................................... viii  

Introduction .................................................................................................................................................. 1  
  The Construct of Self-Regulation ................................................................................................................. 2  
  The Role of Self-Regulation in School Success .......................................................................................... 2  
  The Role of Selective Attention in Self-Regulation and School Success .................................................. 3  
  A Neurophysiological Index of Selective Attention: The P3b Event-Related Potential .............................. 4  
  The Relevance of Individual Differences in P3b Amplitude to Psychopathology, Cognitive Functioning, and Academic Achievement .................................................................................... 6  
  Differences across Socioeconomic Gradients in Self-Regulation and Selective Attention ................... 8  
  The Present Study ...................................................................................................................................... 8  

Methods ....................................................................................................................................................... 10  
  Participants .............................................................................................................................................. 10  
  Teacher-Report Measures ......................................................................................................................... 11  
  Observer-Report Measure ......................................................................................................................... 14  
  Neurophysiological Measures ................................................................................................................. 15  
  Analytic Strategy .................................................................................................................................... 20  

Results .......................................................................................................................................................... 23  
  Univariate and Bivariate Statistics ............................................................................................................ 23  
  Developmental Cascade Models ................................................................................................................ 29  
    Model A: P3b Amplitude Predicting Teacher-Rated Learning Engagement and Academic Performance .......................................................... 29  
    Tests of Indirect Effects in Model A ........................................................................................................ 32  
    Model B: P3b Amplitude Predicting Observer-Rated Task Engagement and Teacher-Rated Academic Performance .................................................................................................................. 34  
    Tests of Indirect Effects in Model B ....................................................................................................... 36  

Discussion .................................................................................................................................................... 39  
  Implications for Science, Practice and Policy ........................................................................................... 41  
  Limitations ............................................................................................................................................... 42  

Conclusion .................................................................................................................................................... 44  

References .................................................................................................................................................... 45
List of Tables

Table 1. Univariate statistics for all analysis variables ................................................................. 23
Table 2. Univariate statistics for teacher-report measures of emotional and behavioral problems .......... 24
Table 3. Pairwise correlations among all analysis variables ................................................................ 26
Table 4. Pairwise correlations of all analysis variables with teacher-report measures of emotional and behavioral problems .................................................................................................................................................. 28
Table 5. Parameter estimates from the best-fitting trimmed model for P3b predicting Learning Engagement & Academic Performance ........................................................................................................................................ 31
Table 6. Estimates of total indirect effects from KG P3b to 1st- & 2nd-grade Learning Engagement & Academic Performance ........................................................................................................................................ 33
Table 7. Parameter estimates from the best-fitting trimmed model for P3b predicting Task Engagement & Academic Performance ........................................................................................................................................ 36
Table 8. Estimates of total indirect effects from KG P3b to 1st-grade Task Engagement & 1st- & 2nd-grade Academic Performance ........................................................................................................................................ 37
List of Figures

Figure 1. Grand-average waveform of correct go-trials at electrode Pz...............................................19
Figure 2. Initial model for P3b predicting Learning Engagement & Academic Performance...............29
Figure 3. Best-fitting trimmed model for P3b predicting Learning Engagement & Academic Performance
........................................................................................................................................................31
Figure 4. Model for tests of indirect effects from P3b to Learning Engagement & Academic Performance
........................................................................................................................................................33
Figure 5. Initial model for P3b predicting Task Engagement & Academic Performance .........................34
Figure 6. Best-fitting trimmed model for P3b predicting Task Engagement & Academic Performance ...35
Figure 7. Model for tests of indirect effects from P3b to Task Engagement & Academic Performance....37
List of Abbreviations

EEG = Electroencephalogram
ERP = Event-related potential
KG = Kindergarten
SEM = Structural equation modeling
Acknowledgements

This work was supported by research funding from the Pennsylvania Department of Health and a pre-doctoral training grant to the author from the Institute of Education Sciences (R305B090007, “Training Interdisciplinary Educational Scientists”).
**Introduction**

Children’s self-regulation of attention, emotion, and behavior is critical for successful school adjustment and optimal academic achievement (Blair & Razza, 2007; Howse, Calkins, Anastopoulos, Keane, & Shelton, 2003; Howse, Lange, Farran, & Boyles, 2003; McClelland et al., 2007; Stipek, Newton, & Chudgar, 2010). Neuroscientists have been making rapid advances in our understanding of the neural mechanisms underlying specific self-regulatory processes, and these neuroscientific advances have provided a theoretical basis for the development of clinical and school-based interventions targeting these neurocognitive systems (Bierman, Nix, Greenberg, Blair, & Domitrovich, 2008; Kerns, Eso, & Thomson, 1999; Posner & Rothbart, 2005). However, basic scientific investigations of how individual differences in these neural mechanisms of self-regulation influence children’s actual behavior and performance in the classroom environment have rarely been conducted.

This paper focuses on the P3b event-related potential as one neurophysiological measure relevant to individual differences in self-regulation. This measure has been widely studied in relation to psychiatric disorders but has received little attention in educational psychology. The amplitude of the P3b reflects neural activity subserving selective attention allocation (Nieuwenhuis, Aston-Jones, & Cohen, 2005; Polich, 2007), a basic neurocognitive process that is instrumental for self-regulatory control (Derryberry & Rothbart, 1997; Rueda, Checa, & Rothbart, 2010; Rueda, Posner, & Rothbart, 2011). The present study tests a hypothesized developmental cascade model (Masten & Cicchetti, 2010) whereby individual differences in neurophysiological functioning subserving selective attention are associated with children’s behavioral self-regulation at school entry, which in turn influences their future academic performance. It is hoped that this study will advance scientific knowledge of the multilevel factors shaping children’s school adjustment and academic performance in the early school years.
The Construct of Self-Regulation

*Self-regulation* refers to the ability to exert goal-directed control over one’s attention, emotions, and behaviors. Independent threads of research across various psychological disciplines, including the study of temperament, developmental neuroscience, educational psychology, and clinical psychology, have converged on the central importance of self-regulatory processes in competent functioning. The construct of self-regulation bears many similarities with the trait that temperament researchers refer to as *effortful control* (Derryberry & Rothbart, 1997; Eisenberg, Smith, & Spinrad, 2011). In developmental neuroscience and neuropsychology, self-regulatory brain processes are often addressed under the rubric of *executive functions* or *cognitive control* (M. W. Cole & Schneider, 2007; Zelazo & Müller, 2010). Educational psychologists also refer to *self-regulated learning*, which additionally incorporates motivational processes and the deployment of specific metacognitive learning strategies (Paris & Paris, 2001). Finally, clinical psychologists refer to *impulsivity* as a multifaceted trait that appears to underlie various externalizing psychopathologies (Hinshaw, 2003). While subtle differences exist between the definitions of each of these terms, they all emphasize self-regulation of attention, emotion, and behavior. Therefore, for the purposes of this paper, I will focus on the substantial overlap that exists across these disciplinary nomenclatures as variations on the broader construct of *self-regulation*.

The Role of Self-Regulation in School Success

Children with weak self-regulatory skills are likely to experience difficulties in the transition to school and throughout their academic career. Teachers generally expect children to display a certain level of self-control in the classroom. Indeed, data from the Early Childhood Longitudinal Study – Kindergarten Class of 1998-99 (ECLS-K), a nationally-representative survey including over 3,000 teachers, revealed that most kindergarten teachers rated self-regulatory behaviors as being more important than academic skills in determining children’s readiness for kindergarten (Lin, Lawrence, & Gorrell, 2003). Behaviors rated as very important or essential by a majority of kindergarten teachers included not being disruptive of the
class, sitting still and paying attention, and finishing tasks. Thus, children with weak self-regulatory abilities will have difficulty meeting the behavioral expectations set by their teachers when they enter school, increasing the likelihood that they will develop conflictual relationships with their teachers (Rudasill & Rimm-Kaufman, 2009). Conflictual relationships with teachers in turn put children at greater risk for escalating levels of aggressive/oppositional behaviors during early elementary school (Silver, Measelle, Armstrong, & Essex, 2005).

Given that engaging and persisting in learning experiences requires self-regulation of attention, emotion, and behavior, poorly regulated children are also inherently at increased risk of poor academic performance due to learning difficulties. Empirical studies have supported an association between self-regulation and achievement in both literacy and mathematics in the early years of schooling (Blair & Razza, 2007; Howse, Calkins, et al., 2003; Howse, Lange, et al., 2003; McClelland et al., 2007; Stipek et al., 2010). Furthermore, deficits in learning-related behaviors including aspects of behavioral self-regulation in kindergarten predict lower rates of growth in literacy and math skills between kindergarten and second grade, resulting in an achievement gap that persists through the sixth grade (McClelland, Acock, & Morrison, 2006). Other studies have found poorly regulated attention and impulse control in kindergarten or earlier to predict academic underachievement into early adulthood even when socio-demographic factors are controlled (McClelland, Acock, Piccinin, Rhea, & Stallings, 2013; Vitaro, Brendgen, Larose, & Tremblay, 2005).

The Role of Selective Attention in Self-Regulation and School Success

Selective attention involves the enhancement of attention to goal-relevant or otherwise salient information and the simultaneous inhibition of attention to goal-irrelevant or distracting information (Dux & Marois, 2009; Neill & Westberry, 1987). This is a core neurocognitive process that is instrumental in the effective regulation of behaviors and emotions in response to challenging situations such as those that confront kindergarten students on a daily basis (Derryberry & Rothbart, 1997; Rueda et al., 2010, 2011). For example, brain imaging studies have revealed that emotional arousal can be modulated by selectively
allocating attention to different features of an arousing stimulus (Ochsner & Gross, 2005). There is also some evidence to suggest that general attention training can reduce emotional reactivity to a stressful task (MacLeod, Rutherford, Campbell, Ebsworthy, & Holker, 2002) and increase behavioral self-regulation in an emotional context (Rueda, Checa, & Cómbita, 2012).

Deficits in selective attention appear to contribute to poor performance on academic tasks such as mathematical problem-solving (Pasolunghi, Cornoldi, & De Liberto, 1999) and reading comprehension (Carretti, Cornoldi, De Beni, & Romanò, 2005), such that poorly-performing students remember less task-relevant information and more task-irrelevant information from verbally presented material. In support of a central role for attention in academic achievement, a recent meta-analysis revealed that attention skills at school entry were a stronger predictor of later literacy and math achievement than were externalizing behavior problems or social skills at school entry (G. J. Duncan et al., 2007). Attentional control has also been found to predict both literacy and math achievement in preschoolers above and beyond the effects of other “executive functions” (Lan, Legare, Ponitz, Li, & Morrison, 2011).

**A Neurophysiological Index of Selective Attention: The P3b Event-Related Potential**

Given the importance of selective attention for children’s self-regulation in the classroom and academic performance, a scientific understanding of the multilevel factors underlying individual differences in selective attention will be key to informing evidence-based intervention strategies and identifying those children who can benefit most from a given intervention. While ample research has documented that the development of competent self-regulation is sensitive to early life experiences such as parenting practices (Kochanska, Murray, & Harlan, 2000; Lengua, Honorado, & Bush, 2007), fewer studies have examined the role of individual differences in neurophysiological functioning.

One neurophysiological measure that has been implicated in selective attention is the P3b event-related potential. The P3b is a late positive peak in the electroencephalogram (EEG) waveform that is typically maximal at centro-parietal electrodes anywhere from 300 to 900 ms following the presentation of a salient, attended visual or auditory stimulus (C. C. Duncan et al., 2009; Linden, 2005; Polich, 2007).
The amplitude of the P3b is sensitive to the selective allocation of attentional resources. For example, stimuli to which participants are instructed to actively attend elicit greater P3b amplitudes than do stimuli that participants are instructed to ignore (Donchin & Cohen, 1967; Hillyard, Hink, Schwent, & Picton, 1973). Additionally, P3b amplitude to target stimuli is diminished as increasingly challenging secondary tasks are added, suggesting that it reflects the amount of attentional resources, in a limited-capacity attentional system, that are available to a particular task (Donchin, 1981; Isreal, Chesney, Wickens, & Donchin, 1980).

The P3b appears to reflect synchronous activity in multiple neocortical regions, including in particular the temporoparietal junction (TPJ) as well as more dorsal posterior areas of the parietal cortex (Linden, 2005). Both of these regions have been implicated in selective attention processes in brain imaging studies. Corbetta and Shulman (2002) conclude that the posterior parietal cortex functions as part of a fronto-parietal network that drives goal-directed attention, whereas the TPJ may function as a “circuit-breaker” that disrupts ongoing attentional processes in order to focus on relevant, unexpected information.

Nieuwenhuis and colleagues (2005) theorized that the synchronous activations in the TPJ and posterior parietal cortices that contribute to the scalp-recorded P3b are produced by phasic activity of the locus coeruleus-norepinephrine system. This phasic activity is hypothesized to enhance the signal-to-noise ratio of neuronal information-processing and thereby to facilitate selective attention (Aston-Jones, Rajkowski, & Cohen, 1999). Interestingly, phasic activity of the locus coeruleus-norepinephrine system exhibits an inverse U-shaped association with tonic activity of this system, similar to that observed for cognitive performance in relation to stress levels. High tonic locus coeruleus activity is associated with low-amplitude phasic responses and poor task performance coinciding with hyperactive and distractible behavior. Thus, low P3b amplitudes may be linked to high tonic levels of noradrenergic signaling which promote relatively indiscriminate neuronal responsivity to irrelevant information (Aston-Jones et al., 1999; Nieuwenhuis et al., 2005). This interpretation of attenuated P3b amplitude as reflecting impaired attentional selectivity clearly has direct parallels with the findings, reported above, that children with poor
reading and math skills tend to recall more irrelevant information than do their more academically skilled counterparts (Carretti et al., 2005; Pasolunghi et al., 1999).

**The Relevance of Individual Differences in P3b Amplitude to Psychopathology, Cognitive Functioning, and Academic Achievement**

The information-processing correlates and probable neural generators of the P3b point towards its utility as a specific neurophysiological measure of selective attention processes that, as reviewed above, are critically important for emotional and behavioral regulation and academic performance. However, most studies of individual differences in P3b amplitude have grouped individuals into psychiatric diagnostic categories rather than examining continua of neurocognitive functioning or non-clinical outcomes such as academic performance. An extensive research literature has implicated low P3b amplitudes in risk for an array of externalizing disorders including ADHD, conduct disorder, and alcohol and other substance dependence disorders (Barry, Johnstone, & Clarke, 2003; Bauer & Hesselbrock, 1999; Iacono & Malone, 2011; Patrick et al., 2006). Patrick and colleagues (2006) proposed that low P3 amplitude is an indicator of a “broad neurobiological vulnerability” underlying the high comorbidity among various externalizing disorders. However, the neurocognitive and behavioral mechanisms of this vulnerability are not yet well understood. Studies using non-clinical samples have the potential to elucidate these mechanisms.

A number of non-clinical studies have suggested that higher P3b amplitudes are associated with better attentional focusing and lower impulsivity/distractibility (Boucher et al., 2010; Harmon-Jones, Barratt, & Wigg, 1997; Määttä et al., 2005; Portin et al., 2000; Russo, De Pascalis, Varriale, & Barratt, 2008). For example, Harmon-Jones et al. (1997) observed that adolescents with higher P3b amplitudes to targets in visual oddball and continuous performance tasks had lower self-rated attentional impulsivity (i.e., less difficulty maintaining focused attention) and lower “non-planning impulsivity” (i.e., a greater tendency to plan for the future). Interestingly, P3b amplitudes were not significantly correlated with self-reported motor impulsivity, suggesting some specificity of the P3b for attentional and cognitive contributors to impulsive behavior. In a sample of middle-aged construction workers, Portin and
colleagues (2000) found that higher P3b amplitudes to targets in an auditory oddball task correlated with faster reaction times on other tasks requiring sustained attention, whereas P3b amplitude did not correlate with reaction times on more cognitively complex tasks or with performance on working memory tasks, again suggesting a specific link between P3b amplitude and attention.

Several studies have also observed positive associations between P3b amplitude and measures of general intelligence in adults (Beauchamp & Stelmack, 2006; De Pascalis, Varriale, & Matteoli, 2008; Russo et al., 2008; Wronka, Kaiser, & Coenen, 2013) and early adolescents (Liu, Xiao, Shi, & Zhao, 2011). Russo and colleagues (2008) observed that P3b amplitude mediated the negative association between impulsivity and IQ, suggesting that highly impulsive individuals may have performed more poorly on intelligence tests due to difficulty inhibiting attentional processing of task-irrelevant information.

Given the association of low P3b amplitudes with greater impulsivity/distractibility and poorer performance on intelligence tests, it stands to reason that P3b amplitude would also have implications for academic performance. However, to date only a small number of studies have examined associations between P3b amplitude and measures of academic performance. Two studies have provided evidence for a positive association between P3b amplitude and concurrent reading achievement among preadolescent and early-adolescent youth (Harmon-Jones et al., 1997; Hillman et al., 2012). Hillman and colleagues (2012) found that this association held above and beyond the effects of socioeconomic status, grade level, and IQ. They also found that the association between P3b amplitude and math achievement was trending in the same direction but was not statistically significant. These studies leave many intriguing research questions unanswered, such as what cognitive or psychosocial processes underlie the association between P3b amplitude and academic performance and at what point in a child’s school experience this association emerges.
Differences across Socioeconomic Gradients in Self-Regulation and Selective Attention

Individual differences in the neurophysiological correlates of self-regulation are particularly important to examine within economically disadvantaged populations, as lower self-regulation may be an important mediator of socioeconomic gaps in academic achievement. Economically disadvantaged children tend to exhibit more self-regulatory behavior problems than do their peers from more advantaged families (Howse, Lange, et al., 2003; Miech, Essex, & Goldsmith, 2001; Noble, Norman, & Farah, 2005; Noble, Tottenham, & Casey, 2005), and these self-regulatory problems are associated with lower reading achievement (Howse, Lange, et al., 2003) and more problematic social relationships with peers and teachers (Miech et al., 2001). Intriguingly, some studies have observed associations between socioeconomic status (SES) and neural mechanisms of attentional control, with individuals of lower SES exhibiting differences in event-related potentials indicative of a decreased ability to ignore distracting and irrelevant information (D’Anguilli et al., 2008; Hackman & Farah, 2008; Stevens et al., 2009). An understanding of the behavioral and academic implications of individual differences in the neural mechanisms of attention may therefore shed light on a neurobehavioral process underlying SES achievement gaps.

The Present Study

There is a scarcity of research on the relevance of P3b amplitude for “real-life” self-regulatory behavior and academic performance in young children just embarking upon their school career. The present study addresses this research gap by testing a hypothesized developmental cascade by which individual differences in kindergarten children’s P3b amplitude may have a lasting impact on their school adjustment and academic performance. Specifically, associations were examined between kindergarten children’s P3b amplitudes in response to target stimuli in a Go/No-Go task, teacher and observer ratings of their behavioral self-regulation, and teacher ratings of their academic performance in kindergarten, 1st, and 2nd grade. It was hypothesized that children with higher P3b amplitudes would exhibit greater self-regulation, which would in turn be associated with higher academic performance. In order to capture the
dynamic, cascading nature of development, longitudinal associations between children’s self-regulation and academic performance from kindergarten to 2nd grade were estimated using cross-lagged autoregressive structural equation models. The structural estimates from these models were used to test the hypothesis that P3b amplitude exerts an indirect effect on children’s academic performance through their behavioral self-regulation (D. A. Cole & Maxwell, 2003).

This study examines these associations within a community-based sample of primarily black and Hispanic children living in low-income urban neighborhoods. Given evidence that both neurophysiological measures of attention processes and behavioral measures of self-control differ across socioeconomic gradients (D’Angiulli, Herdman, Stapells, & Hertzman, 2008; Hackman & Farah, 2009; Howse, Lange, et al., 2003; Miech et al., 2001; Noble, Norman, et al., 2005; Noble, Tottenham, et al., 2005; Stevens, Lauinger, & Neville, 2009), an examination of the developmental processes linking these multilevel factors to academic performance in low-SES samples may yield important policy implications.
Methods

Participants
The sample used for the present analyses consisted of 239 children recruited from elementary schools within a single urban school district in central Pennsylvania. These children were recruited to participate in data collection for a longitudinal randomized controlled trial of a multi-component intervention targeting children with early-onset aggressive behaviors. Of the 239 children in the analysis sample, 107 were rated by their teachers as exhibiting relatively high rates of aggressive behaviors at the beginning of the kindergarten year and were recruited to be in the control group of the randomized controlled trial. The other 132 children were rated by their teachers as exhibiting relatively low rates of aggressive behaviors at the beginning of the kindergarten year and were recruited as a comparison group. Thus, 55% of the children included in the present analyses were in the low-aggression comparison group, and the rest were in the high-aggression control group. An additional 100 high-aggression children who were recruited to be in the treatment group of the randomized controlled trial were not included in the present analyses in order to avoid treatment confounds of longitudinal effects.

The recruitment of high- and low-aggression children was accomplished using screening questionnaires completed by all kindergarten teachers within the participating school district. Teachers completed a 10-item aggressive/oppositional behavior rating scale for each child in their class during the fall of 2008 (Cohort I) and 2009 (Cohort II). Items were drawn from the Authority Acceptance scale of the Teacher Observation of Child Adaptation-Revised (TOCA-R; Werthamer-Larsson, Kellam, & Wheeler, 1991). Teachers rated children on a 6-point Likert scale regarding how often they engage in specific aggressive/oppositional behaviors such as getting in fights or exhibiting cruelty, bullying, or meanness to others. For purposes of the intervention, children were rank-ordered on aggressive behavior within classrooms, and families of children in the upper quartile of aggressive behavior within each
classroom were contacted for recruitment. Of the children targeted for recruitment, 30% were unable to be located or their families refused participation due to the larger demands of the study.

Of the 239 children included in the present analyses, 64% were male. Consistent with the demographics of the region, 69% of the children were African American, 21% were Hispanic or Latino, 9% were Caucasian, and 1% were Asian. The average age at initial screening was 5.65 years ($SD = .36$, range = 5.05 – 6.99). District-wide statistics indicate that these children lived in a community with low socioeconomic resources; 79% of students in the district were classified as low-income (qualifying for free or reduced-price school meals), the majority of households were headed by a single mother (69%), and 79% of parents were estimated to have no more than a high school education. Regional statistics indicate that property crimes were twice as high and violent crimes were 4.5 times as high as comparable statistics for the entire state.

**Teacher-Report Measures**

Teacher reports on participating children were obtained from surveys in the spring of the kindergarten, 1st-grade, and 2nd-grade year for each cohort of students. Teacher response rates were generally high but declined over time. For the control and comparison group sample of 239 children used in the present analyses, 89% of kindergarten teacher surveys, 82% of 1st-grade teacher surveys, and 67% of 2nd-grade teacher surveys were completed. Overall, 132 children (55%) had teacher-report data from all three occasions, 66 children (28%) had teacher-report data from two occasions, 39 children (16%) had teacher-report data from one occasion, and 2 children (<1%) did not have teacher-report data from any occasion. The primary analyses utilized teacher reports of children’s self-regulated engagement in the classroom learning environment (referred to hereafter as “Learning Engagement”) as an indicator of behavioral self-regulation and teacher reports of children’s academic performance as an outcome measure. Additionally, teacher reports of children’s emotional and behavioral problems were examined as ancillary variables of interest. The survey scales composing each of these measures are described below. Univariate statistics for all measures are provided in Tables 1 and 2.
**Learning Engagement.** Teachers responded to 14 items assessing children’s competence in regulating their emotions and behaviors as necessary to engage positively in the classroom learning environment. Representative items included “This child has the self-control necessary to do well in school,” “This child is careful with his or her work,” “This child appears happy and engaged at school,” “This child can get along well enough with other children to succeed in school,” “This child can work independently,” and “This child is able and willing to follow teacher directions.” Responses were provided on a 6-point Likert scale ranging from 1 (“strongly disagree”) to 6 (“strongly agree”). Summary scores were calculated as the mean of the item scores. This measure was adapted from a scale utilized by Bierman et al. (2008) and exhibited very high internal reliability in the current sample (KG: $\alpha=.97$; Gr. 1: $\alpha=.97$; Gr. 2: $\alpha=.98$).

**Academic Performance.** Teachers responded to three questions rating children’s performance relative to other students in their classroom with regard to “reading and pre-literacy skills,” “math skills,” and “overall academic functioning.” Responses were provided on a 5-point Likert scale ranging from 1 (“near the very bottom of your class”) to 5 (“near the very top of your class”). Summary scores were calculated as the mean across these three items. These mean scores exhibited very high internal reliability in the current sample (KG: $\alpha=.99$; Gr. 1: $\alpha=.96$; Gr. 2: $\alpha=.96$).

**Aggressive/Oppositional Behaviors.** This measure was calculated as the mean of standardized subscale scores from two teacher-report scales: the Conduct Problems subscale of the Strength and Difficulties Questionnaire – Teacher Version (SDQ; Goodman, 1997); and an Aggressive/Oppositional Behaviors subscale of a child behavior questionnaire that was developed for the larger study, with some items selected from the Teacher Observation of Classroom Adaptation – Revised (TOCA-R; Werthamer-Larsson et al., 1991). The Conduct Problems subscale consisted of 5 items asking teachers to rate to what degree a statement characterizes each child using a 3-point Likert scale ranging from 0 (“not true”) to 2 (“certainly true”). Items included “often fights with other children or bullies them” and “often lies or...
cheats.” Summary scores were calculated as the mean across all items and exhibited high internal reliability in the current sample (KG: $\alpha=.86$; Gr. 1: $\alpha=.86$; Gr. 2: $\alpha=.87$). On the KG and 1st-grade teacher surveys, the Aggressive/Oppositional Behaviors subscale consisted of 7 items asking teachers to rate how often each child engages in a behavior using a 6-point Likert scale ranging from 1 (“almost never”) to 6 (“almost always”). Items included “yells at others,” “hits, pushes, or shoves,” and “ignores or refuses to obey adults.” On the 2nd-grade teacher surveys, 4 additional items relating to physical aggression were added to the Aggressive/Oppositional Behaviors subscale resulting in a total of 11 items. These additional items included “physically attacks people” and “tries to dominate or bully other children.” Summary scores were calculated as the mean across all items and exhibited very high internal reliability in the current sample (KG: $\alpha=.92$; Gr. 1: $\alpha=.93$; Gr. 2: $\alpha=.97$). The composite score across standardized values on these two subscales also exhibited high internal reliability (KG: $\alpha=.95$; Gr. 1: $\alpha=.95$; Gr. 2: $\alpha=.96$).

**Internalizing Symptoms.** Teachers responded to 6 items regarding how frequently each child exhibits sadness or social withdrawal on a 6-point Likert scale ranging from 1 (“almost never”) to 6 (“almost always”). Items included “sad, unhappy” and “avoids playing with other children.” This scale was used previously in an evaluation of a preschool curricular intervention (http://headstartredi.ssri.psu.edu/). Summary scores were calculated as the mean across all items and exhibited moderate-to-high internal reliability in the current sample (KG: $\alpha=.84$; Gr. 1: $\alpha=.75$; Gr. 2: $\alpha=.80$).

**Hyperactivity/Impulsivity.** This measure consisted of the 8-item Impulsivity-Hyperactivity subscale of the ADHD Rating Scale (DuPaul, 1991). Teachers were asked to report the degree to which a series of statements describe each child on a 4-point Likert scale from 0 (“not at all”) to 3 (“very much”). Items included “has trouble sitting still” and “blurs out answers or opinions inappropriately.” Summary scores were calculated as the mean across all items and exhibited very high internal reliability in the current sample (KG: $\alpha=.94$; Gr. 1: $\alpha=.95$; Gr. 2: $\alpha=.95$).
Inattention. This measure consisted of the 8-item Inattention-Hyperactivity subscale of the ADHD Rating Scale (DuPaul, 1991). Teachers were asked to report the degree to which a series of statements describe each child on a 4-point Likert scale from 0 (“not at all”) to 3 (“very much”). Items included “has trouble staying focused” and “has trouble following directions.” Summary scores were calculated as the mean across all items and exhibited very high internal reliability in the current sample (KG: $\alpha=.96$; Gr. 1: $\alpha=.96$; Gr. 2: $\alpha=.97$).

Observer-Report Measure
Trained research assistants (RAs) administered a battery of tasks designed to directly assess participating children’s cognitive and executive functions in the fall of KG and the spring of 1st grade. In KG, only 5 of the 239 children in the analysis sample did not complete the cognitive assessment; in 1st grade, 35 children were not assessed. Cognitive assessments were also conducted during the spring of 2nd grade but data from these assessments for the low-aggression comparison group are not yet available.

Task Engagement. At the end of the cognitive assessment, RAs retrospectively rated children’s behaviors throughout the assessment period using a 13-item adaptation of the Leiter-R Assessor Report (Roid & Miller, 1997; Smith-Donald, Raver, Hayes, & Richardson, 2007). Items included “pays attention to instructions and demonstrations,” “sustains concentration; willing to try repetitive tasks,” “cooperates; complies with examiner’s requests,” and “modulates and regulates arousal level in self.” In KG, each item was rated on a 4-point scale with item-specific descriptive anchors for each response category. In 1st grade, the items remained the same but the descriptive anchors were removed. Higher scores indicated greater self-regulated engagement. Summary scores were calculated as the mean of all item scores and exhibited high internal reliability in the current sample (KG: $\alpha=.92$; Gr. 1: $\alpha=.93$).
Neurophysiological Measures

Psychophysiological equipment was installed into a recreational vehicle (RV) and driven to each school to conduct assessments. This maximized consistency of the testing environment across the school sites while minimizing burden on parents. In order to reduce apprehension, the RV was decorated with a space theme depicting a familiar cartoon character in an astronaut suit. At the start of each school year all kindergarten classes toured the RV with their teacher to ensure that children were familiar with the vehicle. On the day of the assessment, the tasks were explained to the children verbally, who were then asked to provide verbal assent. Children who asked to return were escorted back promptly, but approached on a separate day and offered the possibility to participate again. Most children who were initially apprehensive eagerly participated on the second day. Only 9 children did not complete a physiological assessment during their KG year. Although physiological assessments were also conducted in 1st and 2nd grade, only the results from the KG assessment are included in the present analyses.

Go/No-Go Task. The task used for the present analyses was an emotional Go/No-Go task modified from a program developed by M. Lewis and J. Stieben (see Lewis, Lamm, Segalowitz, Stieben, & Zelazo, 2006; Stieben et al., 2007). Stimuli consisted of a series of cartoon character images that were designed to be appealing and appropriate for children in kindergarten. Children were told these characters were “alien critters” that they should “zap” by pressing the response box button whenever one appeared (the go condition). The children were then instructed that they were not allowed to zap the same character twice in a row, so they should resist hitting the button if the same critter that they had just seen appeared again (the no-go condition). No-go instructions essentially represented a 1-back working memory condition.

The task consisted of a total of 330 trials, 68% of which were classified as go-trials. Immediately-adjacent stimulus repetitions (no-go condition) occurred on 32% of the trials. A total of 45 different character stimuli were used. Thus, each specific character stimulus had approximately a 2% probability of appearing in each of the 225 go-trials. Success on the task required participants to actively
attend to each go-stimulus in order to compare it to the previous trial and update the working memory trace for comparison with the subsequent trial.

Children completed practice blocks before beginning the testing phase of the task. In order to assess error-related negativity (data not presented here), error rates were targeted to be between 40-60% on no-go trials. Therefore, the initial rate of stimulus presentation (determining the difficulty of the task) was set for the testing phase based on children’s performance during the practice blocks. Additionally, throughout the testing phase, a dynamic computer algorithm tracked no-go error rates; when error rates exceeded 60% the program slowed the trial presentation rate, and when error rates fell below 40% the program speeded the trial presentation rate in order to increase the consistency of task performance across participants. Presentation rate ranged between 435 – 535 ms (M = 515) from the time of response to the next stimulus presentation. Average response time for children in the current analysis sample was 510 ms (range = 332 – 680 ms).

Children were informed that during the game they would be earning points for their performance and that if they got enough points they would win a prize. No specific criterion of point value was associated with earning a prize so children would be motivated to maximize points. Prizes were shown to the child ahead of time and consisted of goody bags filled with small toys, stickers, and other items. Approximately every 10 trials, children were provided with visual feedback informing them of their progress in earning points.

The task was administered in three blocks, with the algorithm for awarding points differing across blocks in order to create experimental conditions of high or low reward. In the first block, the algorithm awarding points strongly favored correct responses and weakly punished incorrect responses, thus resulting in a rapid accumulation of points. In the second block, this algorithm was reversed, resulting in a loss of points regardless of equivalent performance, thus representing a frustrative condition. The final block employed the same high-reward algorithm as the first block and thus all participants ended the game with enough points to win the prize. The entire task lasted approximately 12 minutes.
**EEG Data Recording.** EEG was assessed with a 32-channel elastic stretch BioSemi headcap with the Active Two BioSemi system (BioSemi, Amsterdam, Netherlands). Head circumference was measured to identify cap size. Placement of the Cz electrode was centered at the point of intersection between the line from the nasion to inion and from one temporal mandibular joint to the other. Once the cap was placed, gel was inserted into each electrode receptacle. Two additional electrodes were placed on the left and right mastoids, and four additional facial electrodes were used to measure eye movement. Vertical eye movements were measured from electrodes placed on the infra-orbital ridges centered under the pupils of both eyes and corresponding supra-orbital electrodes embedded within the cap. Horizontal eye movement was measured from electrodes placed approximately 1 cm outside the participants’ right and left outer canthi. Data were recorded at 512 Hz with Actiview Software, v8.0.

**EEG Data Processing.** Data were post-processed using Brain Vision Analyzer 2.0. Voltages were referenced to the average of all electrode sites. EEG data were strongly affected by low frequency power in the delta band, considered typical of young children and conceptualized as a marker of developmental immaturity (Somsen, van’t Klooster, van der Molen, van Leeuwen, & Licht, 1997; Yordanova & Kolev, 2008). In order to reduce the impact of delta frequency noise and maximize the ability to detect individual differences in P3b amplitude in response to the experimental paradigm, some research with young children employs a 1 to 30 Hz bandpass filter (Lewis et al., 2006). Although a high pass filter of 1 Hz is higher than is typically used with adult participants, this setting mitigates the impact of excess delta power while preserving power in the theta frequency, which is thought to contribute to P3b amplitude and is associated with the cognitive ability to encode stimuli (Klimesch, 1999).

Correct responses on go-trials, defined as those for which the child responded to the stimulus between 100 and 1,000 ms after stimulus onset, were segmented from -200 to 1,000 ms relative to stimulus onset. Trials were baseline-corrected to the mean amplitude across the 200 ms prior to stimulus onset and corrections were made for eye blink artifacts using the Gratton and Coles algorithm, as implemented by Brain Vision Analyzer 2.0 (Gratton, Coles, & Donchin, 1983). Any trials with a voltage
step of more than 100 µV between sampling points or a voltage reading outside the range of -75 µV to 75 µV were marked as artifactual and removed from further processing steps.

**Calculation of P3b Amplitude.** P3b amplitudes were calculated across correct go-trials only. Inspection of the grand-averaged waveforms for correct go-trials across all electrodes revealed that, consistent with previous studies (C. C. Duncan et al., 2009; Polich, 2007), voltage increases in the time window associated with the P3b appeared maximal at the midline parietal electrode Pz. The grand-average waveform at the Pz electrode following correct go-trials for children in the current analysis sample is illustrated in Figure 1. P3b amplitude was calculated at this electrode as the mean voltage in the temporal window from 500 to 700 ms post-stimulus, consistent with the location of the peak in the average waveform (see Figure 1) and generally consistent with the temporal characteristics of the P3b observed in previous studies using similarly-aged children (Pfueller et al., 2011; Thomas & Nelson, 1996). Children produced a correct go response for an average of 150 trials across the full task ($SD = 34.4$). Following EEG artifact rejection procedures, an average of 119 ($SD = 38.4$) artifact-free correct go-trials contributed to each child’s mean P3b amplitude, with the number of trials ranging from 22 to 199.¹

¹ P3b amplitude was coded as missing for one child who had only 1 artifact-free correct go-trial. This child’s mean P3b value was greater than 3 SD from the mean.
Figure 1. Grand-average waveform of correct go-trials at electrode Pz

Note. Time 0 marks the stimulus onset. P3b amplitude was calculated as the mean voltage at the Pz electrode in the time window from 500 to 700 ms after stimulus onset.
A total of 211 children had valid P3b amplitude values. In addition to the 9 children who could not be reached for the physiological assessment, EEG data were not collected on 3 children due to braiding of hair that prevented acceptable scalp contact for the EEG. Additionally, P3b amplitude could not be calculated in 13 children due to equipment malfunction (e.g., broken electrodes) or RA error during data collection. Finally, P3b values were recoded to missing for 3 additional children whose values were above or below 3 SD from the sample mean.

**Analytic Strategy**

The frequency distributions of all variables were examined for outliers and departures from normality, and nonlinear transformations were applied where appropriate. The variables of primary interest for the present study are KG P3b amplitude; teacher-reported Learning Engagement in KG, 1st, and 2nd grade; observer-reported Task Engagement in KG and 1st grade; and teacher-reported Academic Performance in all three grades. Bivariate correlations among these variables were examined using pairwise deletion of missing cases. Since roughly half of the sample consists of children exhibiting aggressive behaviors in kindergarten, correlations were also examined between the analysis variables and teacher-reported symptoms of emotional and behavioral disorders.

Cross-lagged autoregressive structural equation models (SEM) were fit to the data in order to test the hypothesized developmental cascade model in which KG P3b amplitude predicts concurrent self-regulation (operationalized using the Learning Engagement and Task Engagement measures), which in turn influences future self-regulation and academic performance over time. In these models, the autoregressive structure accounts for the stability of individual differences in each of these measures over time, while the cross-lagged paths estimate the effects of self-regulation in one year on residualized gains in academic achievement the following year, and vice versa. All SEM models were conducted in LISREL 9.10 using full-information maximum-likelihood (FIML) estimation with manifest variables. FIML estimation takes advantage of all available information provided by the full sample of 239 children, with each case providing information regarding the variables that are not missing for that case (Enders, 2001).
This cross-lagged autoregressive modeling approach has several important advantages over alternate methods such as simple three-variable mediation analyses or correlated latent growth curves. First, the cross-lagged autoregressive model allows for “Granger causal” inferences regarding relations between self-regulation and academic achievement over time (Granger, 1980). For example, self-regulation in one year would be a “Granger cause” of academic performance the following year if it predicted significant variance in next-year academic performance independently of prior academic performance. By controlling for prior levels of academic performance, unmeasured time-invariant confounds that explain the association between self-regulation and academic performance in any year are also controlled. Thus, the autoregressive paths substantially strengthen the ability to make causal claims relative to simple three-variable mediation models (D. A. Cole & Maxwell, 2003; Masten & Cicchetti, 2010). Second, unlike correlated latent growth curve models, which assume that the correlation between change processes is constant over time, autoregressive cross-lagged models permit tests of dynamic causal processes (McArdle, 2009). This is particularly important when examining developmental processes surrounding a major transition period such as school entry, since this period is characterized by sudden qualitative shifts in developmental contexts and behavioral demands that may disturb ongoing developmental processes and set in motion entirely new developmental cascades. Finally, this model allows for longitudinal tests of the total indirect effects of kindergarten P3b amplitude on self-regulation and academic performance in 1st and 2nd grade (D. A. Cole & Maxwell, 2003).

For the present analyses, SEM models were initially specified based on the a priori hypothesized developmental cascade process, as described above and as illustrated in Figures 2 and 5. Separate models were conducted using either teacher-reported Learning Engagement or observer-reported Task Engagement as the measure of behavioral self-regulation. This approach allowed for cross-validation of the estimated structural relations across these two models. Chi-square goodness-of-fit statistics and root mean square error of approximation (RMSEA) values were examined to assess the absolute fit of the models to the raw data. For poorly-fitting models, additional parameters were freed based on examination of the parameter modification indices and theoretical considerations (e.g., not allowing reverse-
chronological predictions). Since the structures of the final models were specified as a result of data-driven decision-making, they should be considered exploratory.

Finally, to explicitly test the hypothesis that children’s behavioral self-regulation mediated the effect of KG P3b amplitude on academic performance in the following two years, the magnitudes of the total indirect effects from KG P3b amplitude to each of the 1st- and 2nd-grade outcomes were calculated.² The statistical significance of each total indirect effect was calculated using the Sobel z-test as implemented by LISREL 9.10.

² Standardized total indirect effects are calculated as the sum of all specific indirect effects. Each specific indirect effect represents one path from the independent variable to the dependent variable and is calculated as the product of all beta coefficients in this path.
Results

Univariate and Bivariate Statistics

Univariate statistics for each analysis variable are reported in Table 1. The skewness and kurtosis values of most of these variables did not reveal substantial deviations from normality. However, the observer-rated Task Engagement measures exhibited moderate negative skew. The distributions of these measures were improved by log-transformation and therefore the transformed values were used in the analyses. Simulation studies have indicated that maximum likelihood parameter estimates and standard errors in SEM are reasonably robust to mild-to-moderate levels of skewness and kurtosis in sample sizes over 100, although chi-square model fit rejection rates are typically inflated with non-normal data (Lei & Lomax, 2005).

Table 1. Univariate statistics for all analysis variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Observed Range</th>
<th>Possible Range</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3b Mean Amplitude (µV)</td>
<td>211</td>
<td>3.40</td>
<td>2.27</td>
<td>-2.7 - 10.0</td>
<td>NA</td>
<td>0.42</td>
<td>0.16</td>
</tr>
<tr>
<td>KG Learning Engagement</td>
<td>211</td>
<td>4.66</td>
<td>1.19</td>
<td>1.1 - 6.0</td>
<td>1.0 - 6.0</td>
<td>-0.75</td>
<td>-0.22</td>
</tr>
<tr>
<td>G1 Learning Engagement</td>
<td>194</td>
<td>4.49</td>
<td>1.28</td>
<td>1.0 - 6.0</td>
<td>1.0 - 6.0</td>
<td>-0.70</td>
<td>-0.31</td>
</tr>
<tr>
<td>G2 Learning Engagement</td>
<td>171</td>
<td>4.42</td>
<td>1.41</td>
<td>1.1 - 6.0</td>
<td>1.0 - 6.0</td>
<td>-0.60</td>
<td>-0.86</td>
</tr>
<tr>
<td>KG Task Engagement</td>
<td>234</td>
<td>3.46</td>
<td>0.55</td>
<td>1.5 - 4.0</td>
<td>1.0 - 4.0</td>
<td>-1.43</td>
<td>1.41</td>
</tr>
<tr>
<td>KG Task Engagement – ln</td>
<td>234</td>
<td>1.00</td>
<td>0.31</td>
<td>0.1 - 1.4</td>
<td>0.0 - 1.4</td>
<td>-0.89</td>
<td>-0.11</td>
</tr>
<tr>
<td>G1 Task Engagement</td>
<td>204</td>
<td>3.32</td>
<td>0.60</td>
<td>1.0 - 4.0</td>
<td>1.0 - 4.0</td>
<td>-1.29</td>
<td>1.91</td>
</tr>
<tr>
<td>G1 Task Engagement – ln</td>
<td>204</td>
<td>0.92</td>
<td>0.33</td>
<td>0.0 - 1.4</td>
<td>0.0 - 1.4</td>
<td>-0.46</td>
<td>-0.35</td>
</tr>
<tr>
<td>KG Academic Performance</td>
<td>207</td>
<td>3.17</td>
<td>1.33</td>
<td>1.0 - 5.0</td>
<td>1.0 - 5.0</td>
<td>-0.12</td>
<td>-1.06</td>
</tr>
<tr>
<td>G1 Academic Performance</td>
<td>194</td>
<td>3.14</td>
<td>1.26</td>
<td>1.0 - 5.0</td>
<td>1.0 - 5.0</td>
<td>-0.14</td>
<td>-1.03</td>
</tr>
<tr>
<td>G2 Academic Performance</td>
<td>173</td>
<td>3.13</td>
<td>1.36</td>
<td>1.0 - 5.0</td>
<td>1.0 - 5.0</td>
<td>-0.15</td>
<td>-1.16</td>
</tr>
</tbody>
</table>

Note. N = Sample Size; M = Mean; SD = Standard Deviation; KG = Kindergarten; G1 = Grade 1; G2 = Grade 2; ln = natural logarithm.

Notably, mean levels of children’s Learning Engagement, Task Engagement, and Academic Performance did not change meaningfully over time. This is to be expected for the Academic Performance measures since teachers were asked to rate participating students’ performance relative to
the other students in their class, and thus children would have to improve relative to their classmates in order for their Academic Performance ratings to increase over time. This was less expected for the Learning Engagement and Task Engagement measures, since they are not explicitly worded to elicit relative rankings. However, since it is well known that children’s absolute levels of self-regulation increase substantially over the early school years, the lack of mean differences in these measures over time indicates that the teachers and observers who were completing these measures were using children’s age-level peers as an implicit reference for what level of behavioral regulation can be expected. Thus, all three of these repeated measures are best suited for measuring individual differences in levels rather than trajectories of change over time.

Table 2 reports univariate statistics for additional teacher-report measures of emotional and behavioral problems that might be expected to correlate with P3b amplitude: Aggressive/oppositional behavior, internalizing symptoms, hyperactivity, and inattention. These variables exhibited mild positive skew, reflecting a low incidence of severe behavior problems despite the fact that nearly half of the sample was selected for high teacher-rated aggressive/oppositional behaviors compared to the other children in their class.

| Table 2. Univariate statistics for teacher-report measures of emotional and behavioral problems |
|----------------------------------|------|-------|---------------|---------------|-------|---------|--------|
|                                  | N    | M     | SD            | Observed Range | Possible Range | Skewness | Kurtosis |
| KG Aggressive/Oppositional Behaviors | 211  | -0.23 | 0.89          | -1.1 - 2.6     | NA              | 1.20    | 0.56    |
| G1 Aggressive/Oppositional Behaviors | 196  | -0.15 | 0.94          | -1.0 - 2.7     | NA              | 1.05    | 0.18    |
| G2 Aggressive/Oppositional Behaviors | 173  | -0.16 | 0.97          | -1.1 - 2.7     | NA              | 1.02    | 0.02    |
| KG Internalizing Symptoms        | 211  | 2.20  | 0.91          | 1.0 - 5.5      | 1.0 - 6.0       | 0.80    | 0.49    |
| G1 Internalizing Symptoms        | 196  | 2.10  | 0.85          | 1.0 - 4.8      | 1.0 - 6.0       | 0.79    | 0.19    |
| G2 Internalizing Symptoms        | 171  | 2.18  | 0.91          | 1.0 - 5.2      | 1.0 - 6.0       | 0.90    | 0.70    |
| KG Hyperactivity/Impulsivity     | 211  | 0.85  | 0.84          | 0.0 - 3.0      | 0.0 - 3.0       | 0.89    | -0.08   |
| G1 Hyperactivity/Impulsivity     | 195  | 0.96  | 0.90          | 0.0 - 3.0      | 0.0 - 3.0       | 0.62    | -0.96   |
| G2 Hyperactivity/Impulsivity     | 172  | 0.80  | 0.89          | 0.0 - 3.0      | 0.0 - 3.0       | 0.92    | -0.41   |
| KG Inattention                   | 211  | 0.96  | 0.96          | 0.0 - 3.0      | 0.0 - 3.0       | 0.66    | -0.68   |
| G1 Inattention                   | 195  | 1.01  | 0.94          | 0.0 - 3.0      | 0.0 - 3.0       | 0.62    | -0.91   |
| G2 Inattention                   | 172  | 0.93  | 0.96          | 0.0 - 3.0      | 0.0 - 3.0       | 0.82    | -0.61   |

Note. N = Sample Size; M = Mean; SD = Standard Deviation; KG = Kindergarten; G1 = Grade 1; G2 = Grade 2.
Pairwise correlations among all analysis variables are presented in Table 3. The repeated measures of teacher-reported Learning Engagement and Academic Performance in KG through 2nd grade showed high stability over time (Learning Engagement range of r’s = .64 - .72; Academic Performance range of r’s = .67 - .76). Observer-rated Task Engagement in KG and 1st grade also showed substantial stability over time, although the coefficient was somewhat lower than those observed for the teacher ratings (r = .41). Additionally, the teacher-rated Learning Engagement and Academic Performance measures were moderately highly intercorrelated within the same school year (range of r’s = .63 - .71). Observer-rated Task Engagement in KG was moderately correlated with concurrent teacher-rated Learning Engagement (r = .57) and Academic Performance (r = .43), but 1st-grade Task Engagement showed more modest concordance with concurrent teacher-rated Learning Engagement (r = .26) and Academic Performance (r = .31). The lower correlations of 1st-grade Task Engagement with related constructs suggest that this measure may have lower construct validity than the KG Task Engagement measure, which could be due in part to the removal of descriptive anchors for each rating category (as discussed in the Methods section), although it could also reflect reduced developmental appropriateness of the measure for 1st-graders. Finally, higher mean P3b amplitudes to target stimuli in the Go/No-Go task were significantly associated with greater teacher-rated Learning Engagement (r = .18) and observer-rated Task Engagement (r = .23) in KG, and with higher teacher-rated Academic Performance (r = .16) in 1st grade. Higher KG P3b amplitudes were also associated with higher teacher-rated Learning Engagement (r = .15) and observer-rated Task Engagement (r = .13) in 1st grade at the trend level (p < .10). Although not reported in the table, P3b amplitude was not correlated with the child’s age in KG (r = -.09, p = .17), nor did it vary according to the child’s gender (t_{208} = -0.92, p = .36) or race/ethnicity (F_{2,208} = 0.09, p = .91).
Table 3. Pairwise correlations among all analysis variables

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. KG P3b Amplitude</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. KG Learning Engagement</td>
<td>0.18**</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(187)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. G1 Learning Engagement</td>
<td>0.15†</td>
<td>0.68**</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(173)</td>
<td>(169)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. G2 Learning Engagement</td>
<td>0.08</td>
<td>0.64**</td>
<td>0.72**</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(151)</td>
<td>(149)</td>
<td>(158)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. KG Task Engagement (ln)</td>
<td>0.23**</td>
<td>0.57**</td>
<td>0.43**</td>
<td>0.49**</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(208)</td>
<td>(206)</td>
<td>(190)</td>
<td>(167)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. G1 Task Engagement (ln)</td>
<td>0.13†</td>
<td>0.32**</td>
<td>0.26**</td>
<td>0.27**</td>
<td>0.41**</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(180)</td>
<td>(178)</td>
<td>(186)</td>
<td>(163)</td>
<td>(200)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. KG Academic Performance</td>
<td>0.07</td>
<td>0.65**</td>
<td>0.41**</td>
<td>0.48**</td>
<td>0.43**</td>
<td>0.27**</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(185)</td>
<td>(206)</td>
<td>(166)</td>
<td>(146)</td>
<td>(202)</td>
<td>(174)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. G1 Academic Performance</td>
<td>0.16*</td>
<td>0.55**</td>
<td>0.63**</td>
<td>0.62**</td>
<td>0.44**</td>
<td>0.31**</td>
<td>0.67**</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>(174)</td>
<td>(170)</td>
<td>(192)</td>
<td>(157)</td>
<td>(190)</td>
<td>(185)</td>
<td>(167)</td>
<td></td>
</tr>
<tr>
<td>9. G2 Academic Performance</td>
<td>0.12</td>
<td>0.50**</td>
<td>0.54**</td>
<td>0.71**</td>
<td>0.40**</td>
<td>0.30**</td>
<td>0.68**</td>
<td>0.76**</td>
</tr>
<tr>
<td></td>
<td>(153)</td>
<td>(151)</td>
<td>(160)</td>
<td>(171)</td>
<td>(169)</td>
<td>(165)</td>
<td>(148)</td>
<td>(159)</td>
</tr>
</tbody>
</table>

Note. Pearson correlation coefficients are reported with pairwise sample sizes in parentheses. KG = Kindergarten; G1 = Grade 1; G2 = Grade 2; ln = natural logarithm. † p < .10; * p < .05; ** p < .01.

Table 4 provides the correlations of all analysis variables with teacher-report measures of other emotional and behavioral problems. Although these correlations are not the primary focus of the present study, they are reported here in order to provide contextual information on the emotional and behavioral correlates of KG P3b amplitude in this unique sample of young children. P3b amplitude was not significantly associated with any of these teacher-report symptom measures at a .05 significance level, although there were trends at a .10 significance level for negative associations of KG P3b amplitude with KG aggressive/oppositional behaviors, 2nd-grade internalizing symptoms, and 1st-grade inattention. Given the non-robustness of these associations, they were not pursued further. It is worth noting that teacher-reported Learning Engagement was highly negatively correlated with each concurrent teacher-reported symptom measure (aggressive/oppositional behaviors: r’s = -.73 to -.76; internalizing symptoms: r’s = -
.60 to -.72; hyperactivity/impulsivity: r’s = -.67 to -.73; inattention: r’s = -.81 to -.84), indicating that the Learning Engagement measure is somewhat redundant with teacher-report measures of a variety of emotional and behavioral problems. However, the fact that the highest correlations were observed with teacher-rated inattention provides some evidence for the construct validity of Learning Engagement as a measure of self-regulatory behaviors that are dependent on selective attention.
Table 4. Pairwise correlations of all analysis variables with teacher-report measures of emotional and behavioral problems

<table>
<thead>
<tr>
<th></th>
<th>KG P3b</th>
<th>KG LE</th>
<th>G1 LE</th>
<th>G2 LE</th>
<th>KG TE</th>
<th>G1 TE</th>
<th>KG AP</th>
<th>G1 AP</th>
<th>G2 AP</th>
</tr>
</thead>
<tbody>
<tr>
<td>KG Aggressive/Oppositional Behaviors</td>
<td>-0.13†</td>
<td>-0.73**</td>
<td>-0.60**</td>
<td>-0.49**</td>
<td>-0.47**</td>
<td>-0.29**</td>
<td>-0.28**</td>
<td>-0.27**</td>
<td>-0.27**</td>
</tr>
<tr>
<td></td>
<td>(187)</td>
<td>(210)</td>
<td>(169)</td>
<td>(149)</td>
<td>(206)</td>
<td>(178)</td>
<td>(170)</td>
<td>(151)</td>
<td></td>
</tr>
<tr>
<td>G1 Aggressive/Oppositional Behaviors</td>
<td>-0.10</td>
<td>-0.44**</td>
<td>-0.76**</td>
<td>-0.54**</td>
<td>-0.28**</td>
<td>-0.23**</td>
<td>-0.09</td>
<td>-0.25**</td>
<td>-0.26**</td>
</tr>
<tr>
<td></td>
<td>(175)</td>
<td>(171)</td>
<td>(194)</td>
<td>(159)</td>
<td>(192)</td>
<td>(187)</td>
<td>(168)</td>
<td>(194)</td>
<td>(161)</td>
</tr>
<tr>
<td>G2 Aggressive/Oppositional Behaviors</td>
<td>-0.05</td>
<td>-0.55**</td>
<td>-0.67**</td>
<td>-0.76**</td>
<td>-0.49**</td>
<td>-0.21**</td>
<td>-0.21*</td>
<td>-0.37**</td>
<td>-0.37**</td>
</tr>
<tr>
<td></td>
<td>(153)</td>
<td>(151)</td>
<td>(160)</td>
<td>(171)</td>
<td>(169)</td>
<td>(165)</td>
<td>(148)</td>
<td>(159)</td>
<td>(173)</td>
</tr>
<tr>
<td>KG Internalizing Symptoms</td>
<td>-0.10</td>
<td>-0.72**</td>
<td>-0.42**</td>
<td>-0.42**</td>
<td>-0.41**</td>
<td>-0.22**</td>
<td>-0.49**</td>
<td>-0.34**</td>
<td>-0.33**</td>
</tr>
<tr>
<td></td>
<td>(187)</td>
<td>(210)</td>
<td>(169)</td>
<td>(149)</td>
<td>(206)</td>
<td>(178)</td>
<td>(206)</td>
<td>(170)</td>
<td>(151)</td>
</tr>
<tr>
<td>G1 Internalizing Symptoms</td>
<td>-0.12</td>
<td>-0.43**</td>
<td>-0.60**</td>
<td>-0.45**</td>
<td>-0.21**</td>
<td>-0.19**</td>
<td>-0.24**</td>
<td>-0.42**</td>
<td>-0.32**</td>
</tr>
<tr>
<td></td>
<td>(175)</td>
<td>(171)</td>
<td>(194)</td>
<td>(159)</td>
<td>(192)</td>
<td>(187)</td>
<td>(168)</td>
<td>(194)</td>
<td>(161)</td>
</tr>
<tr>
<td>G2 Internalizing Symptoms</td>
<td>-0.14†</td>
<td>-0.49**</td>
<td>-0.46**</td>
<td>-0.69**</td>
<td>-0.37**</td>
<td>-0.12</td>
<td>-0.45**</td>
<td>-0.49**</td>
<td>-0.56**</td>
</tr>
<tr>
<td></td>
<td>(151)</td>
<td>(149)</td>
<td>(158)</td>
<td>(171)</td>
<td>(167)</td>
<td>(163)</td>
<td>(146)</td>
<td>(157)</td>
<td>(171)</td>
</tr>
<tr>
<td>KG Hyperactivity/Impulsivity</td>
<td>-0.08</td>
<td>-0.67**</td>
<td>-0.53**</td>
<td>-0.44**</td>
<td>-0.46**</td>
<td>-0.15*</td>
<td>-0.29**</td>
<td>-0.22**</td>
<td>-0.23**</td>
</tr>
<tr>
<td></td>
<td>(187)</td>
<td>(210)</td>
<td>(169)</td>
<td>(150)</td>
<td>(206)</td>
<td>(178)</td>
<td>(206)</td>
<td>(170)</td>
<td>(152)</td>
</tr>
<tr>
<td>G1 Hyperactivity/Impulsivity</td>
<td>-0.10</td>
<td>-0.49**</td>
<td>-0.73**</td>
<td>-0.55**</td>
<td>-0.35**</td>
<td>-0.22**</td>
<td>-0.19*</td>
<td>-0.30**</td>
<td>-0.32**</td>
</tr>
<tr>
<td></td>
<td>(174)</td>
<td>(171)</td>
<td>(193)</td>
<td>(159)</td>
<td>(191)</td>
<td>(186)</td>
<td>(168)</td>
<td>(193)</td>
<td>(161)</td>
</tr>
<tr>
<td>G2 Hyperactivity/Impulsivity</td>
<td>0.03</td>
<td>-0.41**</td>
<td>-0.60**</td>
<td>-0.73**</td>
<td>-0.38**</td>
<td>-0.10</td>
<td>-0.15†</td>
<td>-0.33**</td>
<td>-0.38**</td>
</tr>
<tr>
<td></td>
<td>(152)</td>
<td>(150)</td>
<td>(159)</td>
<td>(170)</td>
<td>(168)</td>
<td>(147)</td>
<td>(158)</td>
<td>(172)</td>
<td></td>
</tr>
<tr>
<td>KG Inattention</td>
<td>-0.09</td>
<td>-0.82**</td>
<td>-0.58**</td>
<td>-0.54**</td>
<td>-0.50**</td>
<td>-0.14†</td>
<td>-0.48**</td>
<td>-0.39**</td>
<td>-0.37**</td>
</tr>
<tr>
<td></td>
<td>(187)</td>
<td>(210)</td>
<td>(169)</td>
<td>(150)</td>
<td>(206)</td>
<td>(178)</td>
<td>(206)</td>
<td>(170)</td>
<td>(152)</td>
</tr>
<tr>
<td>G1 Inattention</td>
<td>-0.13†</td>
<td>-0.57**</td>
<td>-0.81**</td>
<td>-0.66**</td>
<td>-0.43**</td>
<td>-0.32**</td>
<td>-0.32**</td>
<td>-0.50**</td>
<td>-0.48**</td>
</tr>
<tr>
<td></td>
<td>(174)</td>
<td>(171)</td>
<td>(193)</td>
<td>(159)</td>
<td>(191)</td>
<td>(186)</td>
<td>(168)</td>
<td>(193)</td>
<td>(161)</td>
</tr>
<tr>
<td>G2 Inattention</td>
<td>-0.02</td>
<td>-0.49**</td>
<td>-0.65**</td>
<td>-0.84**</td>
<td>-0.42**</td>
<td>-0.19*</td>
<td>-0.30**</td>
<td>-0.49**</td>
<td>-0.57**</td>
</tr>
<tr>
<td></td>
<td>(152)</td>
<td>(150)</td>
<td>(159)</td>
<td>(170)</td>
<td>(168)</td>
<td>(147)</td>
<td>(158)</td>
<td>(172)</td>
<td></td>
</tr>
</tbody>
</table>

Note. Pearson correlation coefficients are reported with pairwise sample sizes in parentheses. KG = Kindergarten; G1 = Grade 1; G2 = Grade 2. LE = Learning Engagement; TE = Task Engagement (log-transformed); AP = Academic Performance. † p < .10; * p < .05; ** p < .01.
**Developmental Cascade Models**

**Model A: P3b Amplitude Predicting Teacher-Rated Learning Engagement and Academic Performance**

The initial hypothesized model which served as a starting point for the SEM analyses linking P3b amplitude to teacher-rated Learning Engagement and Academic Performance over time is illustrated in Figure 2. In this model, P3b amplitude in KG directly predicted concurrent teacher-rated Learning Engagement only. Dynamic causal processes linking Learning Engagement and Academic Performance in KG, 1st, and 2nd grade were modeled using a bivariate first-order autoregressive and unidirectionally cross-lagged structure. In this structure, prior-year Learning Engagement predicted residualized gains in Academic Performance in 1st and 2nd grade, but prior-year Academic Performance did not predict residualized gains in Learning Engagement in 1st and 2nd grade.

**Figure 2. Initial model for P3b predicting Learning Engagement & Academic Performance**

![Diagram](Image)

Chi-Square = 152.1, df = 13, p = 0.00; RMSEA = .219

Note. Standardized path coefficients are presented. Dotted lines represent parameter estimates that are not significant at p < .05. KG = Kindergarten; G1 = Grade 1; G2 = Grade 2.

This model did not fit the data well ($\chi^2 = 162.1$, df = 13, p = 0.00; RMSEA = .219). Not surprisingly, model fit was significantly improved by freely estimating the within-grade residual covariances between Learning Engagement and Academic Performance in the 1st and 2nd grades ($\Delta\chi^2 = 119.5$, $\Delta$df = 2, p = 0.00). Shared rater bias is likely to account for much of this residual within-year covariance, although unmeasured variables predicting between-year variations in both children’s
Learning Engagement and Academic Performance may also contribute. However, this model still did not provide a good absolute fit to the data ($\chi^2 = 42.6$, df = 11, p = 0.00; RMSEA = .110). Therefore, modification indices for the beta parameters were examined to determine which fixed beta parameters could be freed to improve model fit. With the exception of beta parameters representing reverse-chronological predictions, the beta parameter with the largest modification index was freed in an iterative process until no fixed beta parameters had significant modification indices. This procedure resulted in the freeing of three additional beta parameters and produced a model that fit the data well ($\chi^2 = 7.7$, df = 8, p = .46; RMSEA = 0.00). Free beta parameters were added for KG Learning Engagement predicting 2nd-grade Learning Engagement and for KG Academic Performance predicting 2nd-grade Academic Performance, resulting in a 2nd-order autoregressive structure for both of these repeated measures. A significant cross-lagged path was also added from 1st-grade Academic Performance to 2nd-grade Learning Engagement.\(^3\) In contrast, adding a cross-lagged path from KG Academic Performance to 1st-grade Learning Engagement did not improve model fit ($\Delta\chi^2 = 4.5$, $\Delta$df = 4, p = .34). Finally, since the beta parameter for 1st-grade Learning Engagement predicting 2nd-grade Academic Performance was not statistically significant at p < .05, this parameter was fixed to zero in the final model. This best-fitting trimmed model exhibited good absolute fit to the data ($\chi^2 = 8.0$, df = 9, p = .53; RMSEA = 0.00). This model is depicted in Figure 3, and the beta parameter estimates, standard errors, and t-values of each path in this model are reported in Table 5.

---

\(^3\) The beta parameter for KG Learning Engagement predicting 2nd-grade Academic Performance temporarily emerged as statistically significant during the model modification process. However, this parameter became non-significant after the beta parameter for KG Learning Engagement predicting 2nd-grade Learning Engagement was freed. For this reason, the former parameter was fixed to zero in the final model.
Figure 3. Best-fitting trimmed model for P3b predicting Learning Engagement & Academic Performance

![Diagram of the model](image)

Chi-Square = 8.0, df = 9, p = .53; RMSEA = 0.0

Note. Standardized path coefficients are presented. All parameters are statistically significant at p < .05. KG = Kindergarten; G1 = Grade 1; G2 = Grade 2.

Table 5. Parameter estimates from the best-fitting trimmed model for P3b predicting Learning Engagement & Academic Performance

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Outcome</th>
<th>b</th>
<th>SE</th>
<th>t</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>KG P3b</td>
<td>KG LE</td>
<td>0.07</td>
<td>0.03</td>
<td>2.45</td>
<td>0.13*</td>
</tr>
<tr>
<td>KG LE</td>
<td>G1 LE</td>
<td>0.73</td>
<td>0.06</td>
<td>12.09</td>
<td>0.67**</td>
</tr>
<tr>
<td>KG LE</td>
<td>G2 LE</td>
<td>0.31</td>
<td>0.08</td>
<td>3.95</td>
<td>0.26**</td>
</tr>
<tr>
<td>KG LE</td>
<td>G1 AP</td>
<td>0.21</td>
<td>0.07</td>
<td>2.87</td>
<td>0.19**</td>
</tr>
<tr>
<td>G1 LE</td>
<td>G2 LE</td>
<td>0.41</td>
<td>0.08</td>
<td>5.55</td>
<td>0.38**</td>
</tr>
<tr>
<td>KG AP</td>
<td>G1 AP</td>
<td>0.52</td>
<td>0.06</td>
<td>9.03</td>
<td>0.55**</td>
</tr>
<tr>
<td>KG AP</td>
<td>G2 AP</td>
<td>0.28</td>
<td>0.06</td>
<td>4.55</td>
<td>0.27**</td>
</tr>
<tr>
<td>G1 AP</td>
<td>G2 LE</td>
<td>0.23</td>
<td>0.07</td>
<td>3.17</td>
<td>0.21**</td>
</tr>
<tr>
<td>G1 AP</td>
<td>G2 AP</td>
<td>0.63</td>
<td>0.07</td>
<td>9.41</td>
<td>0.59**</td>
</tr>
</tbody>
</table>

Note. KG = Kindergarten; G1 = Grade 1; G2 = Grade 2. LE = Learning Engagement; AP = Academic Performance. b = unstandardized parameter estimate; SE = standard error; t = t-statistic; β = standardized parameter estimate. * p < .10; ** p < .05; *** p < .01.

The results reveal that P3b amplitude in KG significantly predicted teacher-rated KG Learning Engagement. In turn, KG Learning Engagement directly predicted Learning Engagement in 1st and 2nd grade as well as residualized gains in Academic Performance between KG and 1st grade. In contrast, 1st-grade Learning Engagement did not predict residualized gains in Academic Performance in 2nd grade.
Conversely, although KG Academic Performance did not predict residualized gains in Learning Engagement between KG and 1st grade, Academic Performance in 1st grade did significantly predict residualized gains in Learning Engagement in 2nd grade. This pattern is suggestive of a dynamic, non-stationary causal process between these two constructs over the early school years.

In order to explicitly test whether KG P3b amplitude directly predicted gains in Learning Engagement or Academic Performance in 1st or 2nd grade, direct paths from P3b amplitude to each of these outcomes were freed in independent models building upon the final model depicted in Figure 3. None of these direct paths were statistically significant at p < .05. Additionally, the direct path from P3b amplitude to KG Academic Performance was not statistically significant.

Tests of Indirect Effects in Model A
The statistical significance of the total indirect effects from KG P3b amplitude to the 1st- and 2nd-grade Learning Engagement and Academic Performance outcomes was assessed using the Sobel z-test (Sobel, 1982) as implemented by LISREL 9.10. In order to avoid biasing the path coefficients for the indirect effects by fixing small but non-negligible direct effects to zero (Preacher & Hayes, 2008), all direct paths from P3b amplitude to 1st- and 2nd-grade outcomes were freed in the model used for the tests of indirect effects. As expected, this model did not fit the data significantly better than the best-fitting, trimmed model that excluded these direct paths ($\Delta \chi^2 = 4.5$, $\Delta df = 4$, $p = .34$). The model used for the tests of indirect effects is illustrated in Figure 4. The estimates, standard errors, and t-values for each total indirect effect are reported in Table 6.
Figure 4. Model for tests of indirect effects from P3b to Learning Engagement & Academic Performance

![Diagram of the model]

Chi-Square = 3.5, df = 5, p = .62; RMSEA = 0.0

Note. Standardized path coefficients are presented. Dotted lines represent parameter estimates that are not significant at p < .05. KG = Kindergarten; G1 = Grade 1; G2 = Grade 2.

Table 6. Estimates of total indirect effects from KG P3b to 1st- & 2nd-grade Learning Engagement & Academic Performance

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Unstandardized Effect</th>
<th>SE</th>
<th>t</th>
<th>Standardized Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1 Learning Engagement</td>
<td>0.05</td>
<td>0.02</td>
<td>2.43</td>
<td>0.09*</td>
</tr>
<tr>
<td>G2 Learning Engagement</td>
<td>0.06</td>
<td>0.03</td>
<td>2.17</td>
<td>0.09*</td>
</tr>
<tr>
<td>G1 Academic Performance</td>
<td>0.01</td>
<td>0.01</td>
<td>1.82</td>
<td>0.02†</td>
</tr>
<tr>
<td>G2 Academic Performance</td>
<td>0.03</td>
<td>0.02</td>
<td>1.36</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Note. KG = Kindergarten; G1 = Grade 1; G2 = Grade 2. SE = standard error; t = t-statistic. † p < .10; * p < .05; ** p < .01.

There was a statistically significant positive indirect effect of P3b amplitude in KG on 1st-grade Learning Engagement through KG Learning Engagement. Furthermore, there was a statistically significant positive total indirect effect of P3b amplitude on 2nd-grade Learning Engagement. This total indirect effect is the sum of four paths from P3b amplitude to 2nd-grade Learning Engagement: one that includes all the first-order autoregressive coefficients for Learning Engagement between KG and 2nd...
grade, a second that flows from KG Learning Engagement directly to 2nd-grade Learning Engagement via the second-order autoregressive path, a third that includes the (non-significant) direct path from P3b amplitude to 1st-grade Learning Engagement, and a fourth that flows from KG Learning Engagement to 1st-grade Academic Performance to 2nd-grade Learning Engagement. Finally, there was a small, positive indirect effect of KG P3b amplitude on 1st-grade Academic Performance through KG Learning Engagement which was significant at the trend level (p < .10). The total indirect effect of P3b amplitude on 2nd-grade Academic Performance, however, was not significant even at the trend level.

Model B: P3b Amplitude Predicting Observer-Rated Task Engagement and Teacher-Rated Academic Performance

The structure of the initial hypothesized model linking P3b amplitude to observer-rated Task Engagement and teacher-rated Academic Performance is illustrated in Figure 5. This model is structured identically to Model A, except that observer-rated Task Engagement in KG and 1st-grade replace teacher-rated Learning Engagement in these years and the 2nd-grade measure representing the construct of self-regulation is omitted due to incomplete data.

Figure 5. Initial model for P3b predicting Task Engagement & Academic Performance

Chi-Square = 23.6, df = 8, p = 0.00; RMSEA = .090

Note. Standardized path coefficients are presented. Dotted lines represent parameter estimates that are not significant at p < .05. KG = Kindergarten; G1 = Grade 1; G2 = Grade 2.
This model did not provide a good fit to the data ($\chi^2 = 23.6$, df = 8, p = 0.00; RMSEA = .090), and freely estimating the residual covariance between Task Engagement and Academic Performance in 1\textsuperscript{st} grade did not significantly improve model fit ($\Delta\chi^2 = 2.6$, $\Delta$df = 1, p = .11). Based on examination of the beta parameter modification indices, the beta parameter for KG Academic Performance predicting 2\textsuperscript{nd}-grade Academic Performance was freed. The resulting model fit the data well ($\chi^2 = 7.7$, df = 7, p = .36; RMSEA = .021). Freeing the cross-lagged path from KG Academic Performance to 1\textsuperscript{st}-grade Task Engagement did not significantly improve model fit ($\Delta\chi^2 = 1.9$, $\Delta$df = 1, p = .16). Parallel to what was observed for Model A, the beta parameter for 1\textsuperscript{st}-grade Task Engagement predicting 2\textsuperscript{nd}-grade Academic Performance was not statistically significant at p < .05. Therefore, this parameter was fixed to 0 in the final model ($\chi^2 = 9.2$, df = 8, p = .33; RMSEA = .025). This best-fitting, trimmed model is depicted in Figure 6. The estimates, standard errors, and t-values of all beta parameters in this model are reported in Table 7.

Figure 6. Best-fitting trimmed model for P3b predicting Task Engagement & Academic Performance

Chi-Square = 9.2, df = 8, p = .33; RMSEA = .025

Note. Standardized path coefficients are presented. All parameters are statistically significant at p < .05. KG = Kindergarten; G1 = Grade 1; G2 = Grade 2.
Table 7. Parameter estimates from the best-fitting trimmed model for P3b predicting Task Engagement & Academic Performance

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Outcome</th>
<th>b</th>
<th>SE</th>
<th>t</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>KG P3b</td>
<td>KG TE</td>
<td>0.03</td>
<td>0.01</td>
<td>3.13</td>
<td>0.19**</td>
</tr>
<tr>
<td>KG TE</td>
<td>G1 TE</td>
<td>0.43</td>
<td>0.07</td>
<td>6.41</td>
<td>0.41**</td>
</tr>
<tr>
<td>KG TE</td>
<td>G1 AP</td>
<td>0.83</td>
<td>0.24</td>
<td>3.43</td>
<td>0.20**</td>
</tr>
<tr>
<td>KG AP</td>
<td>G1 AP</td>
<td>0.54</td>
<td>0.06</td>
<td>9.42</td>
<td>0.57**</td>
</tr>
<tr>
<td>KG AP</td>
<td>G2 AP</td>
<td>0.30</td>
<td>0.07</td>
<td>4.35</td>
<td>0.29**</td>
</tr>
<tr>
<td>G1 AP</td>
<td>G2 AP</td>
<td>0.61</td>
<td>0.07</td>
<td>8.66</td>
<td>0.57**</td>
</tr>
</tbody>
</table>

Note. Task Engagement measures are log-transformed. KG = Kindergarten; G1 = Grade 1; G2 = Grade 2. TE = Task Engagement; AP = Academic Performance. b = unstandardized parameter estimate; SE = standard error; t = t-statistic; β = standardized parameter estimate. † p < .10; * p < .05; ** p < .01.

According to this model, KG P3b amplitude significantly predicted children’s Task Engagement in KG as rated by a trained observer. KG Task Engagement, in turn, predicted children’s 1st-grade Task Engagement as well as their residualized gain in Academic Performance from KG to 1st grade. In contrast, 1st-grade Task Engagement did not significantly predict children’s residualized gain in Academic Performance in 2nd grade, and KG Academic Performance did not significantly predict residualized gain in Task Engagement in 1st grade. The 2nd-order autoregressive structure of children’s Academic Performance from KG to 2nd grade was essentially identical to that observed in Model A.

Individually freeing the beta parameters for P3b amplitude predicting 1st- and 2nd-grade outcomes revealed that P3b amplitude did not significantly predict gains in Task Engagement or Academic Performance independently of the other predictors in the model. Additionally, as observed in Model A, the direct path from P3b amplitude to KG Academic Performance was not statistically significant.

Tests of Indirect Effects in Model B

In order to accurately estimate the magnitude of the total indirect effects from P3b amplitude to children’s 1st- and 2nd-grade outcomes, direct paths from KG P3b amplitude to 1st-grade Task Engagement and to 1st- and 2nd-grade Academic Performance were freed, producing the model depicted in Figure 7. As expected, this model did not provide significantly better fit to the data than did the best-fitting, trimmed model.
excluding these non-significant direct paths ($\Delta \chi^2 = 2.4, \Delta df = 3, p = .49$). The estimates, standard errors, and t-values for each total indirect effect from P3b amplitude to 1st- and 2nd-grade outcomes are reported in Table 8.

Figure 7. Model for tests of indirect effects from P3b to Task Engagement & Academic Performance

![Model Diagram]

Chi-Square = 6.7, df = 5, p = .24; RMSEA = .038

Note. Standardized path coefficients are presented. Dotted lines represent parameter estimates that are not significant at $p < .05$. KG = Kindergarten; G1 = Grade 1; G2 = Grade 2.

Table 8. Estimates of total indirect effects from KG P3b to 1st-grade Task Engagement & 1st- & 2nd-grade Academic Performance

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Unstandardized Effect</th>
<th>SE</th>
<th>t</th>
<th>Standardized Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1 Task Engagement</td>
<td>0.01</td>
<td>0.00</td>
<td>2.80</td>
<td>0.08**</td>
</tr>
<tr>
<td>G1 Academic Performance</td>
<td>0.02</td>
<td>0.01</td>
<td>2.24</td>
<td>0.04*</td>
</tr>
<tr>
<td>G2 Academic Performance</td>
<td>0.03</td>
<td>0.02</td>
<td>1.49</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Note. Task Engagement measures are log-transformed. KG = Kindergarten; G1 = Grade 1; G2 = Grade 2. SE = standard error; t = t-statistic. † $p < .10$, * $p < .05$, ** $p < .01$. 

37
There was a statistically significant positive indirect effect of KG P3b amplitude on 1st-grade Task Engagement through KG Task Engagement. Additionally, there was a small, statistically significant positive indirect effect of KG P3b amplitude on 1st-grade Academic Performance through KG Task Engagement. The indirect effect of KG P3b amplitude on 2nd-grade Academic Performance did not reach statistical significance.
Discussion

The purpose of this study was to provide insight into how individual differences in neurophysiological functioning underlying selective attention influence young children’s developing self-regulation and academic performance. Specifically, this study examined associations between the amplitude of kindergarten children’s P3b event-related potentials, measured following the presentation of target stimuli in an experimental task, and their self-regulatory behavior and academic performance in KG, 1st, and 2nd grade. Results revealed that higher P3b amplitudes in KG were associated with greater concurrent behavioral self-regulation, which significantly mediated a positive association between KG P3b amplitude and 1st-grade academic performance. This pattern was observed using two independent measures of children’s behavioral self-regulation in KG: teacher-reported Learning Engagement, which reflected children’s self-regulated engagement in the classroom, and observer-reported Task Engagement, which reflected children’s self-regulated engagement during a battery of cognitive assessments.

In both models, this indirect effect was observed above and beyond the stability in children’s academic performance from KG to 1st grade, supporting the interpretation that greater self-regulation in KG caused greater increases in academic performance between KG and 1st grade (according to the logic of Granger causality). The results did not provide support, however, for a significant causal effect of KG or 1st-grade behavioral self-regulation on changes in academic performance in 2nd grade. Additionally, the total indirect effect of KG P3b amplitude on changes in 2nd-grade academic performance did not reach statistical significance.

The observed correlation of P3b amplitude with measures of behavioral self-regulation in KG is consistent with prior research on P3b amplitude and impulsivity/distractibility in older children and adolescents (Boucher et al., 2010; Harmon-Jones et al., 1997; Määttä et al., 2005; Portin et al., 2000; Russo et al., 2008). The current study provides evidence that similar behavioral correlates of P3b amplitude are already discernible in 5- to 7-year-old children. Furthermore, while prior studies used self-
report questionnaires or performance on cognitive assessments to measure impulsivity/distractibility, the present study utilized more ecologically valid measures that capture children’s ability to regulate their attention, emotions, and behaviors in the classroom and while completing a battery of neurocognitive assessments.

The finding of a positive association between KG P3b amplitude and 1st-grade academic performance is also consistent with the small pre-existing literature on the P3b and academic performance (Harmon-Jones et al., 1997; Hillman et al., 2012). These cross-sectional studies found that higher P3b amplitudes were associated with higher concurrent reading achievement in pre- and early-adolescents. However, they provided no empirical evidence regarding how or when P3b amplitude becomes relevant to children’s academic performance. The present study directly addressed this research gap by providing evidence that individual differences in P3b amplitude relate to academic performance as early as 1st grade and that this effect may be mediated by behavioral self-regulation in KG.

It is somewhat surprising that P3b amplitude predicted teacher-rated academic performance in 1st grade but was not associated with concurrent academic performance in KG. One plausible explanation would be that the 1st-grade classroom environment places greater demands on children’s attention and behavioral self-regulation, with the result that the relative learning disadvantage conferred by selective attention deficits increases between KG and 1st grade. Another potential explanation would be that children with poor selective attention fall progressively farther behind their peers over time. However, the explanatory validity of this hypothesis is undermined by the fact that KG P3b amplitude was not significantly associated with 2nd-grade academic performance.

The evidence for a causal path from KG behavioral self-regulation to gains in 1st-grade academic performance is in alignment with an extensive literature suggesting that self-regulatory control at school entry is a critical predictor of children’s academic performance (Blair & Razza, 2007; Howse, Calkins, et al., 2003; Howse, Lange, et al., 2003; McClelland et al., 2007, 2006, 2013; Stipek et al., 2010; Vitaro et al., 2005). However, the lack of a causal path from 1st-grade behavioral self-regulation to gains in 2nd-grade academic performance suggests that the causal processes linking these two constructs over time
may not be stationary. Indeed, in the model utilizing teacher-reported Learning Engagement as the measure of self-regulation, the direction of causality appeared to reverse over time such that 1st-grade academic performance predicted deviations in 2nd-grade self-regulation (this path could not be tested in the model utilizing the observer-reported Task Engagement measures since the 2nd-grade data were not available). This finding should be interpreted with caution since the structural path from 1st-grade academic performance to 2nd-grade self-regulation was identified using an exploratory, data-driven process rather than \textit{a priori} specification. The overall pattern of structural relations suggests that behavioral self-regulation in KG is particularly important for setting the stage for children’s future academic success. By 2nd grade, however, individual differences in academic performance have already become highly self-sustaining (i.e., very strongly predicted by prior academic performance levels), with the result that there is less residual variance to be explained by factors such as prior-year behavioral self-regulation.

It is also interesting to note that, despite the extensive literature linking P3b amplitude to the development of externalizing disorders in older children and adolescents, no significant direct associations (using a significance criterion of .05) were observed in the present sample between KG P3b amplitude and teacher-reported externalizing symptoms in KG, 1st, or 2nd grade. It is possible that the magnitude of the association between P3b amplitude and externalizing disorders increases in older childhood and adolescence.

\textbf{Implications for Science, Practice and Policy}

Since this study used an urban, low-SES sample of children, these findings and their implications are generalizable to a population of great importance for social policy. Support for a relationship between the P3b, as a neurophysiological measure of selective attention, and these children’s behavioral self-regulation and academic performance strengthens the scientific justification for interventions that directly target selective attention skills. The demonstration of an association between P3b amplitude and self-regulation specifically in the early school years is particularly relevant to education practice and policy,
since this is when neurophysiological, behavioral and academic trajectories are most likely to be malleable to intervention. It is possible that children with low P3b amplitudes may be particularly likely to benefit from early interventions targeting selective attention skills. While ERP measurement is highly unlikely to be incorporated into school-based universal screenings due to high cost, difficulty of data interpretation, and a lack of clinically validated cutoff criteria, computer-based tasks in which children’s behavioral performance is closely correlated with individual differences in P3b amplitudes could be used to measure the specific neurophysiological selective-attention processes indexed by the P3b. When reliable measures of neural mechanisms relevant to children’s self-regulation in the classroom have been identified, these measures can be utilized to assess whether interventions have the intended impact on neural functioning and whether changes in neural functioning mediate behavioral changes, thereby simultaneously advancing basic and applied developmental science.

**Limitations**

The results and implications of these analyses must be interpreted with certain limitations in mind. One important limitation is that, since KG P3b amplitude and KG self-regulation are assessed more or less concurrently and without controls for prior levels of self-regulation, the current analyses do not provide statistical evidence for causal effects of KG P3b amplitude on KG self-regulation. It is possible that unmeasured variables, such as SES or parent psychopathology, influence both of these measures. It is also eminently plausible that higher behavioral self-regulation promotes the development of higher P3b amplitudes as a result of experience-dependent plasticity in the neural systems underlying selective attention, or even that reciprocal causal processes between these measures are occurring at a higher time frequency than can be detected from the measurement occasions utilized in the present study (so-called “simultaneous causality”). Indeed, it is hoped that the neural processes underlying P3b amplitude are sensitive to behavioral intervention. Thus, rather than conceptualizing these findings as revealing a unidirectionally causal effect of an immutable biological “trait” on behavior, it may be more appropriate to consider P3b amplitude as a valuable measure of the current state of an individual’s selective attention.
networks. The state of these networks at any given time is undoubtedly influenced by both biological and environmental/behavioral histories, but to the extent that lower P3b amplitude is correlated with poorer concurrent self-regulation it can be seen as a scientifically useful marker of the functioning of a neural system with links to self-regulatory processes.

Another limitation of the present analyses was the use of teacher-reported academic performance as the outcome variable rather than a more objective measure such as standardized test scores. These analyses should be replicated with an objective academic assessment before they can be assumed to reflect the effects of P3b amplitude and self-regulation on children’s true academic competence. However, teachers’ perceptions of students’ academic competence are important variables in their own right, since these perceptions will shape teachers’ behavior towards specific students over the course of the school year. Although this implies a much more complex developmental cascade than was tested in the current analyses, it is at least worth noting that these teacher perceptions may in fact have a stronger impact on students’ future academic achievement than would more objective academic achievement measures such as standardized exams.

The sample used in this study is both a strength and a limitation in terms of generalizability of the findings. This sample represents a hard-to-reach, low-income urban population that is not often included in neurophysiological studies, despite the higher policy relevance of research on these populations. The sample also consists primarily of black and Hispanic children, whereas most neurophysiological studies use primarily Caucasian samples. The racial composition of the sample appropriately reflects the higher concentration of black and Hispanic children in many urban, poor communities in the United States. While the socioeconomic and racial characteristics of the current sample limit the generalizability of the findings to the middle-class Caucasian population that is more commonly studied in neurophysiology, they do support the generalizability of these findings to urban poor communities of high policy relevance.
Conclusion

This study provides evidence that, among urban kindergarten children with low socioeconomic resources, higher P3b amplitudes are associated with superior behavioral self-regulation in the kindergarten classroom, and that this superior self-regulation in kindergarten mediates a positive indirect effect of P3b amplitude on academic performance in 1st grade. These findings fill a significant gap in the research literature by providing support for a specific developmental process that may underlie the association between P3b amplitude and academic achievement and providing evidence that this association emerges as early as 1st grade. Since neurophysiological research suggests that P3b amplitude reflects the activity of neural networks underlying selective attention, these findings build on the scientific justification for interventions targeting young children’s attention skills in order to promote their self-regulation and academic achievement. P3b amplitude, or a closely-linked neurocognitive assessment, may also prove useful as a tool for identifying children to target for specific interventions and for assessing the potential impacts of interventions on the functioning of neural networks subserving selective attention.


