Cascade Model of Executive Functioning, Prosocial Skills,

and Academic Achievement

by

Danielle A. L. Desfossés

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved March 2021 by the
Graduate Supervisory Committee:

Sabina Low, Chair
Marilyn Thompson
Kevin Grimm
Jodi Swanson

ARIZONA STATE UNIVERSITY

May 2021

ABSTRACT

Social Emotional Learning (SEL) programs abound in schools worldwide, adopted in large part on limited and varied evidence that the social/SEL skills acquired in these programs contribute to academic achievement. However, large-scale studies with the most common SEL program in the United States (Second Step®) have yielded no evidence of academic benefits, despite revisions to the Second Step® measure (i.e., DESSA – SSE) to include "skills for learning" (i.e., executive functioning skills). The dearth of academic effects could reflect programmatic or measurement flaws. The purpose of this paper is to explore the latter and unpack the core "inputs" of Second Step® to determine whether the social-emotional or executive functioning components may be differently related to academic achievement. Such questions have important implications for evaluating program theory/logic and for the SEL field more broadly. The current study addresses this broader aim by assessing the longitudinal, bi-directional relationship among Executive Functioning, Prosocial Skills (as a proxy for SEL skills), and academic achievement in Kindergarten and Grade 1 students (N = 3,029) from rural and urban schools (N = 61). Widely utilized curriculum-based measures of reading and math tests were administered directly to students to assess academic achievement, while teachers reported on students' Prosocial Skills using an established measure. A bi-factorial measure of executive functioning was derived from exploratory and confirmatory factor analyses from teacher-reported rating scale data. Results based on autoregressive cross-lagged panel model using accelerated longitudinal design lend some support for a longitudinal bidirectional relationship between the executive functioning components of shifting and emotional regulation (EF 2) and Prosocial Skills. Furthermore, while results support extant research that the executive functioning components of working memory, planning, and problem solving (EF 1) positively predict academic achievement, the executive functioning components of shifting and emotional regulation (EF 2) and Prosocial Skills are not meaningful nor consistent predictors of academic achievement. Implications and limitations are discussed.

DEDICATION

This dissertation is dedicated to my parents, Cathie and Claude Desfossés. I cannot thank you enough for your unwavering support during the writing of this dissertation. While the cost may have seemed beyond reason, your support during this difficult time made it possible for me to achieve my lifelong dream. You have given me the ultimate gift.

ACKNOWLEDGMENTS

Dr. Sabina Low – I could not have written this dissertation without your guidance. Your expertise in crafting the large picture was an enigma at first but in the end, I feel I became an equal partner in the process. While we explored many topics before choosing to focus on children's executive functioning and social emotional learning, I feel that this choice was guided by your understanding of my academic strengths. I have an overwhelming sense of gratitude knowing that were it not for this particular choice of topic, I may not have begun to develop what I hope will become my life's work.

I would also like to thank my committee members Drs. Marilyn Thompson, Kevin Grimm, and Jodi Swanson for your feedback on my dissertation. I remain humbled by your knowledge and expertise in your respective fields.

I would be remiss if I did not also acknowledge Drs. James White and Dan Perlman at the University of British Columbia who undoubtedly provided the foundation to help me transition to and succeed at ASU.

There was not one day in my four years of doctoral coursework at ASU in which I did not feel privileged to be in a classroom. I am grateful to all the professors who made my degree possible. I would especially like to thank – Dr. Carol Martin, who in *Theoretical Issues in Child Development* deepened the way I perceive factors that influence human development more broadly and provided fuel for so many of my papers; Dr. Larry Dumka, who in *Family and Social Policy* taught me the value of a good story; Dr. Gary Ladd, who in *Children's Peer Relationships*, challenged me to think more deeply and critically about child development research within a socio-historical context;

and Dr. Laura Hanish, who in *Prevention in Child Development* provided the knowledge for founding empirically driven programs that bolster human development.

I would also like to thank Dr. Rick Fabes for finding proactive solutions. Dr. Kimberly Updegraff, thank you for having carefully crafted my diverse and comprehensive teaching assistantship portfolio, which inevitably broadened my understanding of family and human development.

It would not have been possible to complete this degree without the support from the Disability Resource Center advisors and directors. I would especially like to thank Angel for being just that.

On a personal note, I would especially like to give my heartfelt gratitude to my peer Shawna who went the extra mile; Michelle, who was always up for a spontaneous adventure; Precious, for being my soul-searching friend; and Lisa, Rob, and Emily, for welcoming me into your family.

Last, I am grateful to my mentors for their devotion to help me succeed. JoAnna, your mentorship provided the means for accelerated personal growth. Joey, you taught me that academic life is the chess board upon which the novice practices the art of strategy, taking no single loss as a defeat, but rather, as an opportunity to improve one's game.

TABLE OF CONTENTS

Page

REVIEW OF LITERATURE (continued)

Page

REVIEW OF LITERATURE (continued)

Page

METHOD (continued)

Page

RESULTS (continued)

Page

RESULTS (continued)

Page

DISCUSSION (continued)

LIST OF TABLES

Table Page

LIST OF FIGURES

INTRODUCTION

Social-emotional learning skills (SEL Skills) are those that enhance children's social and emotional functioning and interactions with others in and out of learning environments. SEL Skills, such as understanding and regulating emotions, problem solving, and prosocial behaviors, have been identified as salient predictors of school success in conjunction with cognitive capabilities (Cambourn, 2002; Denham, 2006; Denham et al., 2012), resulting in increased adoption of school-based interventions to enhance social-emotional learning (SEL). Several evaluations lend support for the success of social skills programs (e.g., Greenberg, & Kusche, 2004; Jones, Brown & Aber, 2011; Low et al., 2015; Marquez et al., 2014; York, 2013). A recent meta-analysis of 213 studies examining the impact of different SEL curricula indicated that such programs are associated with not only significant improvements in students' social-emotional skills but also improvements on end-of-the-year academic achievement (i.e., tests and grades; Durlak et al., 2011). The Second Step® Elementary program (Committee for Children, Seattle, WA) is one of the most widely adopted SEL curricula in schools today.

Second Step® was revised in the last decade to utilize explicit and implicit learning strategies to promote social-emotional as well as targeted executive functioning skills. Executive functioning (EF) skills are cognitive functions that enhance goal-related problem-solving behavior. Second Step® was revised on the premise that the integration of these two domains (SEL Skills and EF) may bolster academic improvements. Despite this, studies to date with Second Step® have yielded inconsistent and weak effects on reading and math achievement (e.g., Low et al., 2019). However, those studies did not distinguish SEL Skills from those grounded in EF theory and literature, limiting knowledge of (a) which components are most impacted by intervention; and (b) which components of EF and SEL are most powerful in explaining academic outcomes.

Despite being closely related, these two constructs come from different theoretical and empirical traditions and for theory and practice, it is important to delineate their respective contribution to academic achievement. The study of EF has been largely undertaken by neurobiologists, neuropsychologists, and to some extent developmentalists (Barkley, 2012a), and has not focused on empirical or conceptual overlap with SEL Skills (Zhou et al., 2012). The investigation of SEL Skills has been almost exclusively examined by prevention scientists, in the education field, whom arguably are more biased toward demonstration of impact than theoretical explication. As a result, some notable gaps exist in our knowledge base.

Using Prosocial Skills as a proxy for SEL Skills, the current study will use existing data from the control group of a longitudinal study of Second Step® in order to facilitate understanding of the relationship between EF and SEL Skills, and their joint role in academic achievement among a diverse population of K to 3rd graders (N = 3,600). Specifically, two main questions will be addressed using a developmental cascade model: (a) what is the longitudinal bi-directional relationship between EF and Prosocial Skills in children?; and (b) what is the additive contribution of examining both EF and Prosocial Skills as predictors of children's academic achievement? Through the shared developmental lens, the ultimate goal of this paper is to test a proposed model that could serve as a framework for additional exploration of the topic useful for both basic and applied science.

REVIEW OF LITERATURE

Executive Functioning

Definitions of Executive Functioning

Over the last thirty years, EF has become a prominent subject of study in neurobiology, neuropsychology, developmental psychology (Barkley, 2012a), and educational and cognitive psychology (Toplak et al., 2013). There is a plethora of definitions within and across the various fields (e.g., see Baggetta & Alexander, 2016; Barkley, 2012; Eslinger, 1996). Based on the work of Luria (1966), Welsh and Pennington (1988) define EF as "the ability to maintain an appropriate problem-solving set for attainment of a future goal" (pp. 201-202). This definition has become rooted in the literature that now views EF as being associated with goal-directed actions (Barkley, 2012b). Other definitions include specific components and areas of the brain associated with EF: "a diversity of hypothesized cognitive processes, including planning, working memory, attention, inhibition, self-monitoring, self-regulation, and initiation carried out by prefrontal areas of the frontal lobes" (Goldstein et al., 2014). Barkley (2012b) defines EF as "acts of self-regulation across time towards future goals" (p. 8), whereby self-regulation refers to six inter-related self-directed activities: attention, inhibition, sensory and motor action (i.e., non-verbal working memory), private speech (i.e., verbal working memory), emotion/ motivation, and play. In short, EF is an umbrella construct that represents interrelated functions concerning "purposeful, goal-directed, problem-solving behavior" (Gioia et al., 2000, p.320).

Organizational Frameworks of Executive Functioning

EF researchers also disagree on an organizing framework (Garon, Bryson, &

Smith, 2008; Huizinga et al., 2006). Three major frameworks are unitary, componential, and unity/diversity. The unitary framework conceptualizes EF as a single component, or one underlying mechanism or ability (Collette et al., 2005) that describes all functions attributed to EF (Huizinga et al., 2006). Many models of cognitive function utilize this framework (e.g., Kimberg & Farah, 1993; Norman & Shallice, 1986). For example, Baddeley's model of working memory (Baddeley, 1986; Baddeley & Hitch, 1974) posits that EF is comprised of working memory, or the central executive, that regulates other subsidiary functions ("subprocesses"). There is some support for the unitary framework for preschoolers (e.g., Wiebe et al., 2011) and children (e.g., Brydges, Reid, Fox, & Anderson, 2012; Brydges, Fox, Reid, & Anderson, 2014).

The second major framework of EF is the componential framework, which considers EF to be comprised of multiple components that can be assessed using factor analysis (Garon et al., 2008). Researchers using the componential framework are further divided among those who believe EF is comprised of either non-correlated components or correlated components. The non-correlated components framework conceptualizes EF as dissociable components (or factors) and is typically used by researchers examining EF in preschoolers based on evidence that executive functions are distinct during early childhood (e.g., Garon et al., 2008), although some (e.g., Peterson & Welsh, 2014) disagree with this conclusion. Alternatively, the correlated componential framework conceptualizes EF as correlated components/factors. Since historically some measures of EF had components that were weakly correlated, Miyake et al.'s (2000) proposed a latent variable approach, which created a factor of EF using multiple EF tests (or scale items). Of all the frameworks of EF, Miyake et al.'s (2000) correlated componential framework is the most commonly used framework (Baggetta & Alexander, 2016; Best, Miller, & Jones, 2009; Peterson & Welsh, 2014). Consequently, the EF research reviewed in this paper almost exclusively refers to the correlated componential framework, unless otherwise indicated.

The componential (correlated or uncorrelated) frameworks sometimes employ a hot/cold analogy to describe the components of EF, with this analogy constituting roughly five percent of research reviewed in Baggetta and Alexander's (2016) meta-analysis. "Cold EF" or the "cool system" (the "know"; Meltcalf & Mitchel, 1999), is typically assessed using some combination of planning, conceptual reasoning, strategic behavior, working memory (Banfield et al., 2004), inhibition, and shifting (Welsh & Peterson, 2014). "Hot EF" or the "hot system" (the "go"; Meltcalf & Mitchel, 1999) describes quick emotional processing and responding of stimulus control involving self-monitoring/self-regulation of emotional processing and response, or emotional regulation. Evidence in support of hot/cool factors in childhood is mixed. Zelazo and Carlson (2012) suggest this may be because unitary and componential frameworks are both used to test the hot/cool framework analogy, typically in children aged 6 or younger, at an age when EF is generally thought to be in an emerging state of formation.

A third emerging framework is Miyake and Friedman (2012) unity/diversity framework that integrates the unitary and the correlated componential frameworks (Garon et al., 2008). This framework posits there is a common EF component that includes inhibition (unity) with specific components of shifting and working memory/updating (diversity). Research that partially lends support to this framework includes that of Monette et al.'s (2015), who performed a confirmatory factor analyses in a sample of normally developing Canadian kindergarten students (N = 272). Somewhat contrary to the he results of their analysis suggest a combined working memory and shifting component and a second component of inhibition.

Major Components of Executive Functioning

As previously explained, most frameworks of EF, with the exception of the unitary framework, classify EF into multiple components. One survey of experts identified 33 components (Eslinger,1996). More recently, Baggetta and Alexander's (2016) meta-analysis (N = 106) using normally developed samples identified 38 components with the four most common components being: (a) *inhibition* (employed in 68% of articles); (b) *working memory/updating* (65% of articles); (c) *shifting* (31% of articles); and (d) *planning* (12% of articles). These same four components are also reported in Best et al.'s (2009) review of the literature of school-aged children. These four components are reviewed below along with a fifth component, that of emotional regulation. This fifth component is reviewed because researchers (typically educators and developmentalists) frequently use childhood rating scales to assess EF. One of the most commonly cited childhood rating scale of EF, the Behavior Rating Inventory of Executive Functioning (BRIEF; Gioia et al., 2000; BRIEF – 2; Gioia et al., 2015), includes scale items representative of emotional regulation.

Table A1 summarizes the definition and alternate name of each component presented below, in addition to listing names of measures by component (cited in-text and in *Appendix C – Commonly Cited Measure of Executive Functioning and Social Emotional Learning Skills*). Measures of EF are numerous with researchers continuing to develop new measures (Baggettta & Alexander, 2016; Jacob & Parkinson, 2015). Types of measures to assess EF differ throughout the lifespan (Ganesalingam et al., 2013), with the two types of measures to assess EF in children being neuropsychological tests or "tasks", often referred to as EF tests (see e.g., Jacob & Parkinson, 2015) and rating scales (see e.g., Baggetta & Alexander, 2016). EF tests are laboratory measures that ask the respondent to follow a task's instructions to measure response time and/or accuracy (Morrison & Grammar, 2016). Rating scales involve a series of questions that measure behavioral manifestations of one or more components of EF via self or other (teacher/parent) reports (Jacob & Parkinson, 2015). Appendix C (*Commonly Cited Measure of Executive Functioning and Social Emotional Learning Skills)* provides an in-depth discussion of some of the most widely used measures of EF, including a critique of EF tests.

Working Memory

While there is no consensus on the definition of working memory (Kimberg et al., 1997; Baddeley, 2012), it is often conceptualized as the ability to "store and manipulate information over brief periods of time" (Alloway et al., 2006, p.1698). Since working memory involves monitoring and updating the relevance of information (Huizinga et al., 2006), an alternative name for working memory is "updating" (Miyake et al., 2000). Numerous tasks are involved in working memory, such as immediate recall, reading comprehension, and problem solving (Cowen, 1998).

While theories of working memory are diverse (Alloway et al., 2006), EF was initially conceptualized as working memory capacity, with a focus on one model in particular (Jurado & Rosselli, 2004); Baddeley and Hitch's (1974) model of the central executive (Baddeley, 1986; Baddeley et al., 1998). This multi-component model specifies a domain-specific storage structure (e.g., verbal, visual, and spatial information) and a domain-general structure, which is the central executive involving the following functions: (a) controls resources and monitors information; and (b) regulates functions, including attention (attentional control) and retrieval/activation of information that is in long-term memory (see also Cowen, 1998). Those ascribing to this model have differing conclusions. For example, evidence suggests verbal and spatial working memory are indistinguishable (for a review, see Kane et al., 2004; Conway et al., 2005), such that some researchers argue that working memory processes are entirely derived from the domain-general structure (Conway et al., 2005).

Inhibition

Many researchers define inhibition, or “self-stopping”, as involving the following three capacities: (a) suppress or disrupt a dominant response; (b) stop behavior towards a goal or an ongoing response that is inefficient; and (c) resist distractions/interference to persist towards the goal (e.g., Barkley, 2012b; Bodnar et al., 2007). In short, inhibition involves controlling or inhibiting behaviors, responses, or thoughts that are dominant or automatic (Baggetta & Alexander, 2016). Alternative names for inhibition in the literature are response inhibition, inhibitory control (Jacob & Parkinson, 2015), and self-restraint (e.g., Barkley, 2012b).

Friedman and Miyake (2004) previously suggested there may be different types of inhibition corresponding to the three capacities listed above, and that these capacities may not necessarily be correlated. However, in proposing their unity/diversity framework, Miyake and Friedman’s (2012) recently suggests inhibition is subsumed under the common EF component, such that it is not a separate component of EF.

Clearly, understanding inhibition still eludes EF researchers (Jurado & Rosselli, 2007), a process that has been underway for nearly a century. Luria (1966) who provided one of the earliest definitions of EF, equated EF with inhibition (Zelazo & Muller, 2002) based on a model of brain development that referenced Vygotsky (1930-1934/1978). Thereafter, the definition of EF began to center on working memory and inhibition based on the works of Diamond (1991) (see Zelazo & Muller, 2002), Roberts and Pennington (1996), and Barkley (1997) (see Brocki & Bohlin, 2004; Thorell & Nyberg, 2008). The definition of EF expanded to include additional components, such as shifting, thereafter.

Shifting

Shifting is "freely moving from one situation, activity, or aspect of a problem to another as the situation demands" (Gioia et al., 2000, p.321). Alternate names including flexibility, cognitive flexibility, attentional set-shifting (Monsell, 2003), attention switching, or task switching (Miyake et al., 2000). An important aspect of switching is being able to adjust behavior according to the situation (e.g., Blair & Ursachi, 2011). An example provided by Monsell (2003) is that of a professor working on an article who switches tasks when the phone rings with an administrative reminder to complete a form and send it to the office prior to resuming work. Dating back to the late 19th Century the study of shifting did not gain attention until the 1990's (Monsell, 1996; 2003), nor prominence until Miyake et al.'s (2000) three component model.

Planning/Problem Solving

Planning is also often cited in the literature as critical to goal-oriented behavior (Best et al., 2009) for it involves the ability to plan actions in advance and to approach a task in an organized, strategic, and efficient manner (Anderson, 2002). While many researchers measure planning as a component of EF, planning is often considered to be a higher-order executive function that first requires the emergence of other major components (Baggetta & Alexander, 2016, Miyake et al., 2000) such as inhibition (see Friedman et al., 2014).

Interest in the components of planning/problem solving is largely driven by the inclusion of representative items on the two most commonly cited rating scales of EF in children (Catale et al., 2015): The Behavior Rating Inventory of Executive Functioning (BRIEF; Gioia, et al., 2000), which has recently been revised (BRIEF – 2; Gioia et al, 2015) and the Childhood Executive Functioning Inventory (CHEXI, Thorell & Nyberg, 2008). Theory also drives interest in the component of planning/problem solving. For example, some componential framework researchers (correlated and non-correlated components) utilize the hot/cold analogy, with planning as a subset of the "cool" component (refer to section *Executive Functioning – Organizational Frameworks of Executive Functioning*). Another example of theory driving interest in this component is Zelazo and Muller's (2010) problem-solving framework, which includes problem representation, planning, execution (of rules/intention), and evaluation error detection and correction.

Emotional Regulation

As there is no consensus on the definition of emotional regulation (Carlson & Wang, 2007; Dodge & Garber, 1991), it is defined here as regulating emotional responses such that the emotion conveyed is appropriate to the situation or stressor (Gioia et al., 2000) for the purpose of facilitating adaptive functioning (Dodge & Garber, 1991). This functioning includes cognitive and social functioning, which some researchers posit are interconnected (e.g., Dodge, 1991). Examples of poor emotional regulation include overreacting to events (e.g., outbursts) and sudden and/or frequent mood changes (Isquith et al., 2013). Numerous theorists across various fields have studied whether and how emotions are regulated by cognitions since the industrial age (Dodge & Garber, 1991; Frijda 1986). Emotional regulation was initially explored in the temperament literature dating back decades, such as to the study of emotional expression in children given favorable and less favorable gifts (Saarni, 1984).

Interest in emotional regulation has primarily been driven by the use of commonly cited EF childhood rating scales and by researchers employing a hot/cold framework (refer to section *Executive Functioning – Organizational Frameworks of Executive Functioning*). The hot system, or the "go" (Meltcalf & Mitchel, 1999) is emotional regulation, which describes quick emotional processing and responding of stimulus control involving self-monitoring/self-regulation of emotional processing and response (Welsh & Peterson, 2014). Emotional regulation is also studied by EF researchers using other frameworks. For example, Hofmann, Schmeichel, and Baddeley (2012) propose components of EF (e.g., working memory) facilitate emotional regulation.

Executive Functioning as a Predictor of Academic Achievement

Elementary school-aged children's reading and math achievement are common foci in the study of EF (Baggetta & Alexander, 2016). However, the relationship between EF and academic achievement is complex because both variables are in a constant state of development during childhood (Bull & Lee, 2014). Furthermore, there are a myriad of ways researchers measure math (e.g., computation, problem solving, etc.) and reading (e.g., spelling, fluency, comprehension). In addition, research on EF as it relates to math abilities tends to focus on only two components (working memory and inhibition), be correlational in nature, and is less often studied in school-aged children (Cragg & Gilmore, 2014; Lemaire & Lecacheur, 2011; Viterbori et al., 2015). EF test-based studies for each of the five aforementioned component of EF (working memory, inhibition, shifting, planning/problem solving, and emotional regulation) are first reviewed, followed by studies that use rating scales. Research using tests and rating scale is reviewed separately due to the very different ways each measures EF (refer to *Appendix C – Commonly Cited Measure of Executive Functioning and Social Emotional Learning Skills*). Namely, rating scales focus on everyday functioning while tests use sequences of movement or knowledge that is intended to mimic every day functioning in a controlled setting. Furthermore, some EF tests only measure one component and other tests measure multiple components, while two commonly cited rating scales (Behavior Rating Inventory of Executive Functioning, Gioia et al., 2000; CHEXI) combine multiple components to minimize the number of factors per scale for analytic purposes.

Research Conducted Using Tests

There is evidence to suggest components of EF develop at different rates in the developing child (Anderson, 2002; Barkley, 2012a; Barkley, 2012b; Best and Miller, 2010; Gioia et al., 2002; Gilberg & Burgess, 2008; Hackman et al., 2015; Hughes et al., 2015; Huizinga et al., 2006; Macdonald et al., 2014; McCann, 2010; Miller et al., 2013; Romaine & Reynolds, 2005; Welsh, 2001). Consequently, a more accurate understanding of the relationship between EF and academic achievement is provided herein by reviewing single components of EF as a predictor of academic achievement (i.e., working memory, shifting, planning/ problem solving, emotional regulation) in addition to research examining EF as a multi-component construct predicting academic achievement (i.e., working memory and inhibition; working memory, inhibition, and shifting).

Working Memory. Literature reviews of studies assessing children of various ages suggest working memory accounts for unique variance in math achievement (e.g., written calculations, verbal calculations, and word problems) (Cragg & Gilmore, 2014; Friso-Van den Bos et al., 2013). Working memory has long since been established as contributing to speed and span of math computations as measured by arithmetic (Adams & Hitch, 1997). In De Smedt et al.'s (2009) one-year longitudinal study of Grade 1 Belgian children (N = 77), hierarchical regression analyses suggested working memory, assessed as a composite score via seven EF tests (e.g., Counting Span, Case, Kurland, & Goldberg, 1982; Listening span test, van der Sluis et al., 2005; Working Memory Test Battery for Children, Pickering & Gathercole, 2001) to tap Baddeley's working memory model significantly predicted Grade 2 math achievement (e.g., arithmetic and measurement), controlling for Grade 1 math achievement and intellectual ability ($F_{(4,69)}$ = 19.82, $p < .001$, $R^2 = .65$).

Working memory is seldom examined as a predictor of reading ability, but more so in reading comprehension, in which it has predominantly been studied in populations of school aged children with reading disabilities (see Swanson et al., 2004; Swanson et al., 2009). In one study that examined predictors of math and reading ability, Geary (2011) employed multilevel modeling analysis to a longitudinal study of American children (N = 177) from Grade 1 through Grade 5. Working memory was measured via all nine subscales of the Working Memory Test Battery for Children (WMTB–C; Pickering & Gathercole, 2001), and math and reading achievement were measured via the

Numerical Operations and Word Reading subtests of the Wechsler Individual Achievement Test–II (WIAT-II; Wechsler, 2001). Results suggest the relationship of working memory (central executive) to math achievement increased with age. Working memory was also a significant predictor of reading scores for Grade 1, but not Grade 2.

Working Memory and Inhibition. Bull and Lee (2014) caution that studies that measure the relationship between inhibition and math achievement find a significant relationship only because these studies included working memory in their analyses. However, some studies have partial out the variance of working memory and are thus reported here.

Bull and Scerif (2001) conducted a series of regression analyses based on a sample (N = 93) of Scottish children (M_{age} = 7.3). Inhibition, as measured by the Wisconsin Card Sorting Task (Heaton, Chelune, Talley, Kay, & Curtiss,1993) and Stroop Task (Stroop, 1935) accounted for 21% of the variance in group math test including single and multi-digit addition and subtraction accounted ($p < .001$). Similarly, running the model with only a working memory latent variable, comprised of Counting Span (Case et al., 1982) and Dual Task Performance (Baddeley et al., 1997) accounted for 19% of the variance in math scores ($p < .001$). When both inhibition and working memory were included in the model, these variables accounted for 33% of the variance in math scores ($p < .001$).

Another study of inhibition and academic achievement that partials out the effect of working memory was conducted on British school children (N = 51; M_{age} = 11.8) that also used the Stroop Task (Stroop, 1935) to measure inhibition reported moderate and significant correlations with composite scores of latent variables in English ($r_{ab, c}$ = .31, p

< .05) and standardized math scores ($r_{ab,\,c}$ = .36, $p < .05$), holding working memory constant (St. Clair-Thompson & Gathercole, 2006). Thus, there is some limited evidence to suggest inhibition may affect academic achievement, controlling for working memory.

Shifting. Research on the relationship between shifting and math achievement is limited (Cragg & Gilmore, 2014; Lemaire & Lecacheur, 2011). Yeniad et al., 2013) conducted two meta-analyses of concurrent and longitudinal studies on shifting and academic achievement. Children in these studies were from normally developing populations and were primarily aged three to ten. Research articles on shifting were restricted to performance-based tasks of shifting (e.g., EF tests), as it relates to math (k = 18, N = 2330) and reading (k = 16, N = 2266). Shifting was significantly and equally associated with math achievement (r = .26, 95% CI = .15–.35) and reading achievement (r = .21, 95% CI =.11–.31). While Yeniad et al. (2013) controlled for IQ, Bull and Lee (2014) point out that this review did not control for working memory in their analyses and as such, suggest these findings be interpreted with caution.

Working Memory, Inhibition, and Shifting. In Neuenschwander et al. (2012) longitudinal sample of Swiss, mostly German-speaking Grade 2 children of diverse socioeconomic background (N = 459; M_{T1} = 7.4 yrs.), EF was assessed using a single latent construct based on three EF tests representing three components of EF: (a) working memory (Backward Color Recall Task, Schmid et al., 2008)

(b) inhibition (e.g., aforementioned Fruit Stroop Test variation (Archibald & Kerns, 1999); and (c) shifting (i.e., Cognitive Flexibility task, Zimmermann et al., 2004). Two latent variables were created for the academic achievement: (a) standardized achievement tests for math (3 arithmetic tests) and literacy (2 reading tests and 1 writing test); and (b) grades composite of math, reading, and writing. The model was a good fit to the data ($\chi^{(48)} = 106.3$, $p < .001$, CFI = .98, RMSEA = 0.052). The paths from EF to math and literacy were significant, although stronger evidence was found based on the standardize regression coefficients for the path between EF and standardized tests versus grades ($ß_{tests} = .88$, $p < .001$; $ß_{grades} = .40$, $p < .001$). Similarly, in another model tested by Neuenschwander et al. (2012) using the above noted standardized achievement tests for math and literacy, the model was also a good fit to the data ($\chi^{(48)} = 106.3$, $p < .001$, CFI = .98, RMSEA = 0.052). The paths from EF to math and literacy were significant ($ß_{math} = .93$, $p < .001$; $ß_{literacy} = .80$, $p < .001$). Literacy and math were reported to predict EF equally well since constraining the paths to be equal did not worsen model fit ($\chi^2_{(1)} = .6$, $p = .61$).

In a longitudinal study of mostly low income, ethnically diverse pre-kindergarten ($M_{age} = 4.49$) American children ($N = 164$), Sasser et al. (2015) created a composite measure EF (1-factor) using three EF tasks: working memory was assessed using the Backwards Word Span (Davis & Pratt, 1996); working memory and inhibition was assessed using the Peg Tapping task (Diamond & Taylor, 1996); and shifting and inhibition was assessed using the Dimensional Change Card Sort (DCCS; Frye et al., 1995). Separate HLM with repeated assessment nested within children were conducted on each of the four academic outcome measures for children's scores in Grade 3. These outcome measures included: (a) two reading achievement measures: The Letter-Word Identification subtest of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, & Mather, 2001) and the Test of Word Reading Efficiency (Torgensen et al., 1999); (b) math achievement of number comprehension, adding, and subtracting

(Woodcock et al., 2001); and (c) teacher-rated academic functioning on reading, writing, and math (3-items; Academic Performance Rating Scale, DuPaul et al., 1991). Results suggest statically significant differences in the intercept (ß = 0.28, *SE* = 0.08, $p < 0.01$) and slope (ß = 0.24, *SE* = 0.09, $p < 0.01$) in math achievement only. Thus, a one standard deviation increase in a child's pre-kindergarten EF score resulted in a 0.24 increase in the Grade 3 math achievement score. There was no observable inter-individual variability in Grade 3 intercept nor in the rate of growth (since kindergarten) for either of the reading measures. With regards to academic functioning, results suggest significant differences in the intercept only (ß = 0.23, *SE* = 0.08, $p < 0.01$).

Longitudinal studies have also been conducted examining the relationship between EF and academic achievement, with a 2-factor representation of EF assessed using working memory and *combining* inhibition-shifting. Two such studies are summarized below due to the similarities in their findings. In Van der Ven et al.'s (2012) study of Grade 1 Dutch children (*N* = 211) assessed across four waves (spanning 6 months each) over two years, initially nine EF tests were used to create latent variables for the following components: working memory (i.e., Digit Span Backwards and Odd One Out tasks, Alloway, 2007; Keep Track, van der Sluis, de Jong, & van der Leij, 2007); inhibition (i.e., two author created tests and the Animal Stroop, Wright, Waterman et al., 2003); and shifting [i.e., two author created tests and the Children Coloured Trail Test (Llorente et al., 2003)]. Math achievement was measured using standardized tests (largely comprised of addition, subtraction, multiplication, division, and measuring) with good psychometric properties employed by the Dutch School Board. Since the 3-factor model did not fit the data, a series of models were tested using various combination of two-factor models were tested (e.g., inhibition and shifting/ working memory; inhibition and working memory/ shifting; shifting and updating/ inhibition). The model with a latent variable of working memory and one that combined inhibition with shifting was selected but the authors could not generate a growth curve model that fit the data sufficiently. Thus, a growth curve model with only working memory and math scores was a good fit to the data ($\chi^{(114)} = 117.6$, $p = .39$, CFI = 1.0, RMSEA = 0.01), suggesting children with high working memory scores had high math scores.

A second longitudinal study of interest on the relationship between EF and academic achievement was conducted by Lee et al.'s (2012) on six-year-old children (N = 163) from low to middle socioeconomic backgrounds in Singapore. Eleven EF tests in total were used to create latent variables of: working memory, three tests (i.e., Listening recall, Mister X, and Pictoral Updating from Alloway, 2007); and inhibition-shifting, eight tasks from three tests, all of which were modified (Flanker task, Fan, McCandliss et al., 2006; and Number-letter task, Miyake et al., 2000). Results suggested that working memory at Time 1 but not the combined latent variable inhibition-shifting at Time 1, predicted math achievement at Time 2, as measured by the Numerical Operations subscale of the WIAT-II (Wechsler, 2001) ($\chi^{(321)} = 446.82$, CFI = 0.92, RMSEA = 0.05). Thus, in these two studies (Lee et al., 2012; Van der Ven et al., 2012), only working memory and not the latent inhibition-shifting variable predicted math achievement.

Planning/Problem Solving. EF's component of planning/problem solving has seldom been examined in relation to academic achievement with longitudinal data (Baggetta & Alexander, 2016; Best et al., 2009). Friedman et al.'s (2014) longitudinal study of American children (N = 1,364) created latent reading and math scores based on the Woodcock–Johnson Psycho-Educational Battery– Revised (Woodcock & Johnson, 1989– 1990; Woodcock & Mather, 1989) and planning based on The Tower of Hanoi Test (Piaget, 1974/1976). Results of latent growth curve analysis suggest that for Grade 1 children, a 1 SD in the planning intercept was associated with a 0.26 SD increase in reading score intercepts (ß = 0.97, $p < .001$) and .47 increase in math score intercept (ß = 1.39, $p < .001$). Similarly, for Grade 3 children, a 1 SD in the planning intercept was associated with a 0.16 SD increase on the reading score intercept (ß = 0.43) and 0.37 increase in math score intercept (ß = 0.84) intercept.

Emotional Regulation. The emotional regulation literature is a study unto itself with the development of a measure for school-aged children (Shields & Cicchetti, 1997) extending this research to academic achievement (e.g., Graziano et al., 2007). Within the context of EF, EF test-based research has emphasized the "cold" EF components (e.g., working memory, shifting) over the "hot" EF component of emotional regulation (Barkley, 2012a; Peterson & Welsh, 2014) (refer to section *Organizational Frameworks of Executive Functioning*). Consequently, research on emotional regulation and early academic achievement is scarce (Graziano et al., 2007). Research that employs a hot/cold EF framework in school-aged samples tends to focus on the transition to school using preschool measures of EF adapted from self-regulation measures (e.g., Brock, Rimm-Kaufman et al., 2009; Willoughby et al., 2011), which do not strictly measure EF as previously discussed. Other frameworks of EF that include emotional regulation such as Barkley's (2012a; 2012b) have not extended this research to academic achievement. Additionally, research using the aforementioned Children's Gambling Task (Kerr & Zelazo, 2004) does not appear to have been extended to the study of school-aged children's academic achievement, although Buelow and Barnhart (2017) recently examined math performance using the Iowa Gambling Task in a sample of adults. While research that examines emotional regulation using tests of executive functioning is scare, our knowledge of this topic can be extended by studies conducted using rating scales.

Research Conducted Using Rating Scales

The Behavior Rating Inventory of Executive Functioning (BRIEF; Gioia et al., 2000; BRIEF – 2, Gioia et al., 2015) is the most commonly used rating scales to assess EF in children (Catale et al., 2015). Longitudinal data using the BRIEF (Gioia et al., 2000) and academic achievement in school-aged children largely reports on correlations. For example, in a sample of predominantly low income minority American Grade 5 children (N = 91), Waber et al. (2006) reported moderately strong correlations between standard-based math tests and the BRIEF's BRI subscale (e.g., inhibition, shifting, and emotional control) (r = - .47, p < .001); and the MI subscale (e.g., initiation, working memory, plan/organize, monitoring, and organization of materials) (r = - .61, p < .001). Similarly, Dekker (2017) evaluated EF in a sample of Dutch children aged 6-8 years (N = 84) at two time points. EF was measured using the parent and teacher versions of the Behavior Rating Inventory of Executive Functioning (BRIEF; Gioia et al., 2000) based on a one-factor solution that combined working memory/planning, shifting, and inhibition). Math was measured using Dutch standardized and norm-referenced (e.g., addition/subtraction). While small significant correlations were reported between math (T1) and working memory/planning (T2) for the parent (r = -.28, p < .01) and teacher (r = -.23, p < .05) reports, nonsignificant correlations were reported for shifting (rs = -.11, -.13) and inhibition (rs = .08, .04) for the parent and teacher reports, respectively.

The Childhood Executive Functioning Inventory (CHEXI; Thorell & Nyberg, 2008) is the second most commonly used rating scales to assess EF in children (Catale et al., 2015). Longitudinal investigations using the CHEXI (Thorell & Nyberg, 2008) are needed, particularly in academic achievement, to fully utilize the scale's utility (Thorell & Catale, 2014). In a sample of Swedish kindergarten children (M_{age} = 76 months), Thorell and Nyberg (2008) administered the parent ($n = 113$) and teacher ($n = 105$) versions. Significant medium sized correlations were reported controlling for EF tests (Go No-Go; word span task of the Digit Span subtest of WISC-III) for: (a) math and working memory/planning for the parent ($r = -.29$, $p < .001$) and teacher ($r = -.42$, $p < .001$) versions; (b) language and working memory/planning for the parent ($r = -.41$, $p < .001$) and teacher ($r = -.46$, $p < .001$) versions; and (c) smaller but mostly significant correlations were reported for both math and language with inhibition. In Thorell et al's. (2013) sample of children aged 6 to 11 in Spain ($n = 219$), Sweden ($n = 141$), Iran ($n = 49$), and China ($n = 72$), significant medium to large correlations were reported between working memory/planning and both math and reading using both versions in Sweden and Spain and using the teacher version in China and Iran (r s = -.28 to -.59, ps < .01), controlling for age. Smaller, significant correlations were reported between inhibition and math and reading using both versions in all countries (rs = -.19 to -.48, ps < .05).

Social Emotional Learning (SEL) Skills

Definitions of SEL Skills

Historically, child development programs have focused on promoting positive skills or preventing negative outcomes (Catalano, Hawkins, Berglund, Pollard, & Arthur

2002), with a melding of the two frameworks into programs targeting the development of social emotional learning (SEL) skills (Catalano et al., 2004). SEL refers to the process by which skills, attitudes, and knowledge are acquired to "understand and manage emotions, set and achieve positive goals, feel and show empathy for others, establish and maintain positive relationships, and make responsible decisions" (CASEL, 2012, p.4). Thus, SEL *Skills* include managing emotions, caring about others and developing positive relationships, making good decisions, avoiding negative behavior, and behaving ethnically and responsibly (Elias & Moceri, 2012). SEL Skills could be defined more broadly as those that enhance children's social and emotional competencies, functioning, and interactions with others in and out of learning environments (Dr. Stephen Elliott, *personal communications*, April 25, 2017).

Daniel Goleman's 1995 New York Times bestseller Emotional Intelligence became instrumental in the dissemination of SEL (Hoffman, 2009; Elias & Moceri, 2012), and to the creation of an organization called Collaborative for Academic, Social, and Emotional Learning (CASEL), an international leader in the field of SEL (Durlak et al., 2011; Zins et al., 2004a). It was around this time that SEL programs emerged. While the initial focus among researchers was on how SEL programs were preventing negative outcomes, the promotion of academic outcomes became prominent in the research (Goleman, 2004), with these programs flourishing largely in school systems in the US (CASEL, 2013) and around the world (Torrente et al., 2015). SEL programs target the development of a variety of SEL Skills (Rimm-Kaufer & Hulleman, 2015), akin to previous historical efforts of developing social skills (Gresham & Elliott, 2008; Gresham et al., 2018). Standards for SEL programs have been established by organizations such as

CASEL. While the application of free-standing SEL standards in preschools have been established in 50 states, the development and application of state-wide elementary school SEL standards is still emerging in elementary schools (Dusenbury & Weissberg, 2016). Currently, only four states are using these standards (i.e., Illinois, Kansas, West Virginia, and Pennsylvania), although several other states are in the process of developed free standing standards and guidelines or have developed these standards for certain grades.

One SEL Skills area is prosocial behavior. Interest in prosocial research spanned from the 1970's until the early 1990's and only recently has the topic re-emerged (Eisenberg & Spinrad, 2014). Prosocial behavior has been studied from various theoretical perspectives (e.g., evolutionary, cognitive-developmental, moral-socialization) (Carlo, 2014), but developmentalists often define it as a "voluntary behavior intended to benefit another" (Eisenberg et al., 2006, p.646; also see Eisenberg & Spinrad, 2014; Holmgren et al.,1998). Prosocial behavior requires social cognition, motivation to act, and anticipation of a reward (Chakroff & Young, 2014). Research has focused on environmental influences of prosocial behavior, such as the various motivations propelling one to act prosocially (e.g., Carlo et al., 2003; Eisenberg, 1986; Penner et al., 2005). Biological influences of prosocial behavior emerged in the research in the 1990's (Penner et al., 2005) with some evidence suggesting prosocial behavior is roughly half genetic influence and half environmental influence (e.g., Scourfield et al., 2004).

Organizational Framework of SEL Skills

Various frameworks of SEL Skills have been proposed (e.g., CASEL, 2012; Common Core State Standards; Rimm-Kaufer & Hulleman, 2015; Payton et al., 2000; Zins et al., 2004b). The most widely accepted framework is that proposed by CASEL

(2012), which consists of five interrelated sets of cognitive, affective, and behavioral competencies. These competencies, also referred to as CASEL Core-5 (e.g., Zins et al., 2004b), are based on the work of Elias et al. (1997) (Weissberg et al., 2015) and a five-competency framework initially proposed by Zins et al. (2004b). CASEL Core-5 competencies are as follows: (a) *Self-Awareness* of one's thoughts and emotions and how those relate to behavior; (b) *Self-Management* of thoughts, emotions, and behaviors as they relate to regulating stress and impulses and sustaining motivation and goals; (c) *Social Awareness* includes having an awareness of available social support as well as understanding social and ethical mores for behavior (e.g., taking the perspective of and empathizing others who have different backgrounds and cultures); (d) *Relationship Skills* includes communication, cooperation, negotiation, resisting social pressure, and proving/seeking social support; and (e) *Responsible Decision* making with regards to oneself and others (CASEL, 2012).

Major Dimension of SEL Skills

As evident from the description of CASEL's 5-core competencies of SEL above, Social Emotional Learning (SEL) is a diverse area in the literature comprising many competencies (Catalano et al., 2004). One such competency, that of prosocial behavior, is a suitable proxy for SEL Skills because many SEL programs specifically target the promotion of prosocial behavior (CASEL, 2012), with results of improvement in participants' prosocial behaviors and attitudes (Durlak et al., 2011). Additionally, prosocial behavior is a suitable proxy for SEL Skills because it comprises a major domain of one of the most cited measures of SEL Skills, the Social Skills Information System Rating Scale (SSIS – RS; Gresham & Elliott, 2008).

Prosocial Behavior

Similar to the definition of EF discussed earlier, prosocial behavior is an umbrella construct that represents interrelated behaviors such as sharing, helping, comforting, and cooperating (Carlo, 2014; Weir & Duveen, 1981). While most researchers conceptualize these interrelated behaviors as correlates of prosocial behavior, which is measured as a unidimensional construct (Crowe et al., 2015; Padilla-Walker & Carlo, 2014), many notable researchers have argued that prosocial behavior is a multidimensional construct or a construct that requires a multi-level approach (e.g., Batson, 1998; Carlo & Randall, 2002; Eisenberg & Fabes,1998; Eisenberg et al., 2006; Padilla-Walker & Carlo, 2014; Penner et al., 2005).

Considering prosocial behavior as a multidimensional construct calls for the use of global measures that assess prosocial behaviors across contexts and/or motives. For example, Carlo and Randall (2002) proposed a six-factor model of the motivations that explain a person's decision to engage in prosocial behavior (e.g., public, anonymous, dire, emotional, compliant, altruism). This typology is measured using the Prosocial Tendencies Measure (PTM) devised by the authors. Richaud et al. (2012) tested this model under the direction of Carlo using a translated version of the PTM on a sample of Argentinian children (N = 472, $Mage$ = 12.4, SD = 1.6). Results of a confirmatory factor analyses suggests a four-factor solution (i.e., altruistic, public, anonymous, and responsive) was a good fit to the data (χ^2 (183) = 461.35, p < . 001, CFI = .93, RMSEA = .057).

In addition to considering prosocial behavior across context and/or motives, Eisenberg et al. (2006) suggest a comprehensive understanding of the construct should also include, for example, individual factors such as cognitive capacity to assess the situation, make inferences of the other's needs, emotional and/or biological (e.g., genetics) capability of reacting, and ability (e.g., knowledge or willingness) to take action. As such, the aforementioned interrelated behaviors such as sharing, helping, comforting, and cooperating are only one aspect towards an understanding of prosocial behavior.

While to date, no such comprehensive measure has been developed, as mentioned above, prosocial behavior comprises a major domain of one of the most widely used measures of SEL Skills: the Social Skills Improvement System Rating Scales (SSIS; Gresham & Elliott, 2008). Its offspring, the Social Skills Information System SEL Edition Rating Form (SSIS SEL – RF; Gresham & Elliott, 2017) is considered to be a measure of prosocial behavior (Gresham et al., 2011; Kettler et al., 2011), and was created to align with the CASEL Core-5 (Elliott et al., 2017; Gresham et al., 2018), as described further in Appendix C (*Commonly Cited Measure of Executive Functioning and Social Emotional Learning Skills*).

Prosocial Skills as a Predictor of Academic Achievement

Meta-analytic studies suggest there is a large body of evidence that SEL Skills are related to school outcomes (e.g., Durlak et al., 2011; Taylor et al., 2017). As previously discussed, research on prosocial behavior (i.e., Prosocial Skills) is used herein as a proxy for SEL Skills. Prosocial Skills is a component of many elementary school intervention programs such as the Peace Builders Program (Flannery et al., 2003) and the Social Skills Improvement System Classwide Intervention Program (SSIS-CIP; Elliott & Gresham, 2007). Efficacious programs such as these increase Prosocial Skills (e.g., Diperna et al., 2015; DiPerna et al., 2017). Additionally, the Collaborative for Academic, Social, and Emotional Learning (CASEL; 2013) found support for a significant, positive relationship between Prosocial Skills and academic achievement for elementary school intervention programs such as the Positive Action Program (Flay et al., 2001; Flay & Allred, 2003), 4R's Program (Jones et al., 2011), and Resolving Conflict Creatively Program (Brown et al., 2004).

Outside of the context of intervention programs, evidence also suggests a relationship between prosocial behavior and academic achievement. Earlier research tended to be correlational in nature (Eisenberg et al., 2006) or use other traditional statistical techniques. For example, using a diverse sample of 3rd and 4th Grade students in the Eastern US, teacher-rated social skills including prosocial behavior (cooperation, assertion, responsibility, self-control, and empathy) at Time 1 was moderately correlated with Time 1 standardized reading ($r = .49$, $p < .001$) and math ($r = .31$, $p < .01$) scores as well as in a six month follow up at Time 2 [reading ($r = .40$, $p < .001$), math ($r = .37$, $p < .01$)] (Malecki & Elliott, 2002). Stronger evidence is provided by Wentzel and Caldwell's (1997) data using largely Caucasian Grade 6 ($N = 404$) children followed over three years reported a medium sized correlation ($r = .59$, $p < .001$) between prosocial behavior based on a teacher rated two-item measure (helping and considerate) and grade point average for 6th graders, with a similar correlation ($r = .59$, $p < .001$) reported in a follow-up two years later.

The prosocial subscale of the SDQ (Goodman, 1997) was positively correlated with standardized reading ($r = .38$, $p < .001$) and math ($r = .28$, $p < .001$) tests in the aforementioned cross-sectional study of 364 English native-speaking children in the UK ages 8-11 (Adams et al., 1999). However, Adams and colleagues (1999) reported that a hierarchical linear regression suggested prosociality contributed only a small amount of unique variance on reading ($r^2 = .034$, $p < .001$) and no significant variance on arithmetic ($r^2 = .002$, *ns*). When interpreting findings, it is important to consider problems associated with traditional statistics. For example, Romano et al. (2010) demonstrated that running Ordinary Least Squares regression analyses without proper missing data handing techniques can lead to false-negative findings regarding the association between various behaviors, including prosociality, and academic achievement. Regression analyses using an iterative multiple imputation approach for missing data (i.e., Markov chain Monte Carlo method) on a previously published dataset consisting of largely English speaking (72.8%), socioeconomically diverse Canadian children (N = 1,521) suggested prosociality in Kindergarten predicts marginally higher levels of reading (ß = .11, SE = 0.03, $p < .001$) and math skills (ß = .07, SE = 0.03, $p < .01$) in Grade 3.

Longitudinal analyses examining prosocial behavior and academic achievement with more advanced statistics is limited. In a study by Miles and Stipek (2006), a small sample of ethnically diverse sample of low-income American children in kindergarten or Grade 1 were followed through to Grade 5 (N =102). Measures administered included The Child Behavior Scale (Ladd & Profilet, 1996), a teacher-report measure of prosociality (4 items: help, empathetic, shows concern, offers help, α = .86-.89) and standardized scores based on reading and comprehension tests for literacy achievement (α = .74-.82). Standardized parameter estimates from path analyses conducted using "various model fit statistics in LISREL" (p.111) suggest prosocial behavior in Grade 1 predicted literacy achievement in Grade 3, while holding Grade 1 literacy achievement constant (ß = .19, $p < .001$). However, prosocial behavior in Grade 3 did not predicted literacy achievement in Grade 5, while holding Grade 3 literacy achievement constant (ß = -.02, *ns*).

In a second example of research examining the relationship between prosocial behavior and academic achievement results of structural equation modeling analyses by Caprara, Barbaranelli, Pastorelli, Bandura, and Zimbardo (2000) using a sample of economically diverse Italian children ($N = 294$) suggests third grade prosociality (M_{age} = 8.5) predicts eighth grade academic achievement ($b = 0.52$, $p < .05$), with similar results in a subsample ($n = 100$) that controlled for Time 1 academic achievement ($r = .57$). The authors used a multi-informant approach of self-reports (α = .78), peer-reports (α = .78), and teacher reports (α = .89) created by the authors to devise a latent prosocial variable based on a 10-item scale (help, kind, share, cooperate) and academic achievement was a latent variable based on six teacher ratings of children's achievement, which corresponds to Italian education standards of assessment.

A third longitudinal study employing advanced statistics concerns a secondary data analysis by Caemmerer and Keith (2015) of ethnically diverse American children (N = 7,802) followed longitudinally over six waves from kindergarten to Grade 8. A latent achievement variable was created for each grade based on reading (e.g., letter and vocabulary knowledge, reading comprehension, α = .87-.96) and math (e.g., operations, measurement, algebra, α = .91-.95) (Tourangeau et al., 2009). A composite social skill construct was created based on an adaptation of the teacher-rated Social Skills Rating Scale (SSRS; Gresham & Elliott, 1990), comprising of items from the problem behavior and social skills subscales (α = .79 - .89). While this variable is not prosocial behavior per say, because there is overlap between SSRS and prosocial behavior (refer to *Appendix C – Commonly Cited Measure of Executive Functioning and Social Emotional Learning Skills*) and the moderate stability of the social skills measure over time (b = .46 to .71, SE = .02–.04), the complexity of the analysis warrants a discussion of the results. A latent variable longitudinal structural equation panel model was tested first by comparing nested models using Sattora–Bentler corrected $\chi 2$ using a cross-validation approach. The baseline model was a good fit to the data (χ^2 = 749.661, df = 204, RMSEA = .028, CFI = .976, SRMR = .049). Results suggested small, statically significant paths between social skills and academic achievement for some of the waves. Specifically, kindergarten social skills (wave 2) predicted Grade 1 achievement (wave 3) (b = .911, β = .058, SE = .021, p = .005), and Grade 5 social skills predicted Grade 8 achievement (b = 1.099, β = .066, SE = .017, p = .001), controlling for gender, SES, and verbal ability.

In conclusion, there is some mixed evidence on the relationship between prosocial behavior and academic achievement, with earlier correlational studies suggesting a moderate relationship, but longitudinal studies using more advanced statistical methods yielding mixed results.

Empirical Relationship Between Executive Functioning and SEL Skills

Research Overview

Researchers have long been interested in the empirical relationship between EF and SEL Skills. For example, in 1985 Hartup proposed: “the cognitive functions most closely linked to social relationships are the ‘executive regulators’ – the planning, monitoring, and outcome-checking skills involved in problem-solving” (p.75). One decade later, SEL began to emerge as a concept to describe how individuals learn social skills in the context of relationships. At that time in Goleman's (1995) proposed in his New York Times bestselling book, *Emotional Intelligence*, that SEL affects certain components of EF (i.e., emotional impulses (regulation) and working memory) with the underlying premise that SEL programs are lessons in emotional intelligence (Goleman, 2004).

Testable models of the relationship between EF and SEL Skills emerged after Goleman's (1995) publication (e.g., see Riggs et al., 2006). Various proxies were used to examine SEL Skills. Initially, researchers examined the relationship between EF and social emotional deficits or difficulties such as ADHD (for a review see Riggs et al., 2006). While research in this area continues (e.g., Huang-Pollock et al., 2009), thereafter, researchers focused on the relationship between EF and social competence (Bierman & Erath, 2006).

Social competence is a proxy for SEL Skills that forms the basis of many school based SEL interventions (Elias et al., 1997). The study of social competence concerns diverse skills such as cooperative play, communication, emotional understanding, self-regulation, aggression control, and social problem-solving skills (Bierman & Erath, 2006). Note that social competence's skill of *social* problem solving different from EF's planning/problem solving component (refer to section *Major Components of Executive Functioning*), which concerns the ability to plan *actions* in advance and to approach a task in an organized, strategic, and efficient manner (Anderson, 2002). As such, construct overlap is not being identified as a concern herein in models that include EF and social competence's skill of social problem solving. With regards to the social competence skill of self-regulation, however, as described in the following section (see *Limitations*), because EF and self-regulation do share some similarities (e.g., construct overlap), research that examines the relationship between EF and social competency using self-regulation may have validity issues.

Direction of the Relationship Between Executive Functioning and SEL Skills

The research examining the relationship between EF and SEL Skills, regardless of the proxy for SEL Skills (e.g., social competency as a proxy) tends to be unidirectional, emphasizing EF as a predictor of SEL Skills (Greenberg et al., 2004; Stitcher et al., 2016). Select studies from the social competency literature are summarized below.

In a two-year randomized control trial of the SEL intervention PATHS® (Greenberg & Kusché, 1993), Bierman et al. (2008b) examined the relationship between EF and SEL Skills in a sample of 356 ethnically diverse American kindergarten children (T1 M_{age} = 4.49 yrs.). SEL Skills (i.e., sharing, helping, empathy, resolving conflict) was measured by the teacher version of the Social Competence Scale (Conduct Problems Prevention Research Group, 1995, α =.88). Two measures of EF were administered at T1: a) working memory/inhibition (Peg Tapping Task, Diamond & Taylor, 1996); and b) working memory, inhibition, and shifting (Dimensional Card Sorting, Frye et al., 1995). Both measures of EF significantly predicted SEL Skills at T2 (ß = .20, $p < .001$, ß = .17, $p < .01$) and T3 controlling for verbal IQ and age. Contrarily, Sasser et al.'s (2015) study of mostly low income, ethnically diverse American children (N = 164), pre-kindergarten EF skills (latent score of working memory, inhibition, and shifting using various EF tests) did not directly predict growth in children's SEL Skills (ß = 0.12, *SE* = 0.09, *n.s.*), using the same measure, the Social Competence Scale (Conduct Problems Prevention Research Group, 2003).

Data from a variety of pediatric medical populations (e.g., Kiley-Brabec & Sobina, 2006; Wolfe et al., 2013) with diverse age ranges (i.e., 4 – 18 years) suggests SEL Skills using the most commonly cited measures of SEL Skills, the SSIS – RS (Gresham & Elliott, 2008) and its parent scale the SSRS (Gresham & Elliott, 1990) significantly predicts EF, measured using the BRIEF (Gioia et al., 2000). With regards to research applicable to school-aged children, Hensler et al. (2013) study of children with Sickle Cell disease (N = 20, M_{age} = 11.25) using hierarchical linear regression analysis suggested the proportion of variance accounted for by EF using a rating scale (i.e., BRIEF- GEC, or Global Score, Gioia et al., 2000) an EF tests (i.e., NEPSY-II Animal Sorting subtest, Korkman, Kirk, & Kemp, 2007) and IQ (i.e., WASI; Wechsler, 1999), accounted for 65% of the variance in social skills, as measured using the SSIS – RS (Gresham & Elliot, 2008).

There is some new evidence emerging in the preschool literature that suggests that the relationship between EF and SEL Skills may be bidirectional such that the development of one influences the development of the other (Morigushi, 2014). This emerging research was not reviewed herein because it concerns a different population, which in turn has different ways of measuring the EF and SEL Skills constructs. For example, EF is measured differently in the preschool literature due to components that are not thought to emerge until elementary-school age (e.g., planning/problem solving). The preschool literature is also more focused on social interaction versus SEL Skills. Both EF and SEL Skills in the preschool literature also use Theory of Mind, which is seldom applied to the elementary school literature of these constructs (Hughes & Devine, 2015).

While the bidirectional relationship between EF and SEL Skills is beginning to be explored empirically in the preschool literature, examining the bidirectional relationship between EF and SEL Skills in elementary school aged children would extend our understanding of the relationship between these two constructs. This bidirectional relationship is further explored in the study presented at the end of this literature review (refer to section *Current Study*) and forms the basis for Research Question 1 for the purpose of filling a gap in this literature.

Empirical Relationship Among Executive Functioning, SEL Skills, and Academic Achievement

The empirical relationship among EF, SEL Skills, and academic achievement is important because the Common Core State Standards (CCSS) deems the measurement of children's cognitive and SEL Skills a priority in relation to improving children's educational outcomes. CCSS is a major initiative to which most US states subscribe (Common Core Standards Initiative, 2012) and is in line with CASEL Core-5 (Dusenbury et al., 2015). The study of this relationship seems a natural extension, then, of the considerable research that examines the relationship between EF and academic achievement (refer to section *Executive Functioning as a Predictor of Academic Achievement*) and SEL Skills and academic achievement (refer to section *Social Emotional Learning Skill of Prosocial Behavior Predicts Academic Achievement*).

Due to the plethora of ways the constructs of EF, SEL Skills, and academic achievement can be measured, research that measures these three constructs in the same model is even more diverse. As such, to sift through this body of literature I established two main search criteria, both relating to the academic achievement construct. First, proxies for academic achievement were limited to reading or math, which are common foci in the study of EF (Baggetta & Alexander, 2016). These two proxies are in keeping with the current study's measures (refer to section *Method*). As such, studies examining the relationship among these constructs but that use more specific proxies for academic achievement such as verbal formation/expression (e.g., see Riggs et al., 2006) or broader proxies, such as intelligence (e.g., see Fishbein et al., 2016), were excluded from the search. Second, the research had to measure academic achievement as an outcome variable rather than, for example as a predictor (e.g., see Weiland & Yoshika, 2014). In the research that emerged that adhered to these two main criteria, I identified two central limitations with the applicability of the research to the study of the empirical relationship among these constructs.

The first limitation that became apparent in my review of the empirical relationship among these constructs is that it is often discussed in the school readiness literature which typically focuses on preschoolers or Kindergarteners (e.g., Bierman et al., 2008; Bierman, Torres, Domitrovich, Welsh, & Gest, 2009; Mann, 2012). This literature cannot be applied to the study of EF in elementary school-aged children because of the way EF and academic achievement are often measured in this population. EF is measured as a unitary construct in the preschool literature (e.g., Wiebe et al., 2011), since major components of EF are not thought to be clearly distinguishable from each other until early childhood (Best et al., 2009). Similarly, academic achievement is not measured in the same way in preschool and Kindergarten as it is in early elementary school aged children. More global measures of achievement are often used (e.g., cognition and general knowledge) rather than specific measures such as reading and math (e.g., see Flook et al., 2015).

The second limitation that became apparent in my review of the empirical relationship among these constructs concerns the conceptual overlap between EF and SEL Skills that can occur. This overlap is discussed in detail in the following section (refer to section *Limitations*). This research may include proxies of SEL Skills such as learning related social skills/behavior (e.g., Brock et al., 2009; Nesbitt et al., 2015; Lee, 2016), effortful control (e.g., Blair & Razza, 2007), or self-regulation (e.g., Blair & Dennis, 2010; Morrison et al., 2010).

Based on the research that was excluded due to the two limitations identified above (i.e., school readiness literature and conceptual overlap between EF and SEL Skills) and the inclusion of the search criteria identifying academic achievement as an outcome variable with reading and math as proxies, research applicable to the empirical relationship among these three constructs was extremely limited. Applicable research examined only one component of EF. Rhoades et al. (2011) measured the mediating role of attention to the study of SEL Skills and achievement in school-aged children. However, attention is not one of the five components discussed herein and was not a measure in my study (refer to section *Method*). Upon first glance, other research appeared relevant. For example, Ladd et al.'s (2014) study of collaborative skills and academic achievement in school-aged children measured planning/problem solving as a collaborative skill. However, this research was not measuring – nor was it attempting to measure – an EF component using comparable tests or rating subscales. In conclusion, no studies were identified that examined the relationship among EF, SEL Skills, and academic achievement that fit my search criteria.

Because academic achievement outcomes matter to schools and the Common Core State Standards and because EF and SEL Skills are often promoted in interventions used by schools, further investigation of the relationship among these constructs is warranted. Thus, this relationship is further explored in the study presented at the end of this literature review (refer to section *Current Study*) and forms the basis for Research Question 2 for the purpose of filling a gap in this literature.

Conceptual Relationship Between Executive Functioning and SEL Skills

Background

One challenge for prevention scientists and those in the SEL field is attention to areas of conceptual and measurement overlap with EF. While the concepts of EF and SEL Skills are highly related, their underlying measurement must be distinct to minimize issues of discriminant validity. This in turn minimizes the probability of multicollinearity when EF and SEL Skills are assessed in the same model. Ensuring that the measurement of these constructs remains distinct will in turn clarify the contribution of each construct to academic achievement, and ultimately, to assessments of SEL intervention program validity.

Eisenberg and Zhou (2016) initiated a discussion on the conceptual overlap between EF and SEL Skills such as self-regulation and effortful control (Eisenberg & Zhou, 2016). Given that researchers use either effortful control or EF to study childhood self-regulation (Blair & Razza, 2007; Liew, 2012; Zhou et al., 2012), it is surprising that the literature rarely discusses the conceptual overlap among these constructs. I propose to extend this discussion in two ways herein: (a) by further elaborating on the relationship among EF, temperamentally-based self-regulation, and effortful control, followed by providing case examples from this literature as well as from the learning-related skills literature; and (b) by highlighting overlap among items from some of the most commonly cited childhood EF and SEL Skills rating scales.

This discussion applies to future rating scale developers of EF and SEL Skills in addition to researchers who may otherwise unintentionally use both constructs in their model without fully understanding their potential for multicollinearity. In so doing, the broader implication of this section is to better equip researchers in the field of SEL Skills to be able to unpack program effects on each constructs' contribution to academic achievement.

Executive Functioning, Self-Regulation, and Effortful Control

EF, which is sometimes conceptualized as self-regulation, most notably in the writing and childhood EF rating scale of Dr. Russel Barkley (e.g., BDEFS – CA; Barkley, 2012b) is conceptually distinct from temperamentally-based self-regulation. While there is little consensus on the definition of (temperamentally based) self-regulation (Bronson, 2000; Eisenberg & Zhou, 2016; McClelland & Cameron, 2012), it is often referred to as the ability to control attention and behavior (Derryberry & Rothbart, 1988) in relation to cognitive, emotional, and social demands (Posner & Rothbart, 2000). Temperamentally based self-regulation is often referred to as effortful control (Rothbart & Rueda, 2005; also see Eisenberg et al, 2011; Eisenberg, 2012). Most studies of self-regulation are in fact measuring effortful control (Blair & Razza, 2007). Consistent with Rothbart's research refining the dimensions of effortful control (Derryberry & Rothbart, 1988; Evans & Rothbart, 2007; Rothbart, 2011; Rothbart & Rueda, 2005; Rothbart et al., 2001; Rothbart et al., 1994), Eisenberg and Zhou (2016) state effortful control is comprised of the following dimensions: (a) attention focusing (i.e., maintaining focus on tasks); (b) attention shifting (i.e., directing attention towards a more appropriate stimulus); (c) inhibitory control (i.e., plan or suppress responses during instruction or during situations of ambiguity); and (d) activation control (i.e., perform and action despite tendency to do otherwise).

Similarities

First, EF and effortful control are similar in the sense that they both have components/dimensions of *emotional regulation*, which are similarly measured. It is important to clarify that some literature refers to self-regulation as emotion-related self-regulation, sometimes referred to as *emotional regulation*. This term, coined by Eisenberg and Spinrad (2004), distinguishes between sources of regulation (internal or external) and processes (conscious or unconscious), with voluntary processes referred to as what Rothbart and colleagues label effortful control (Eisenberg & Zhou, 2016). The EF literature also frequently measures emotion-related self-regulation, or *emotional regulation*, via rating scales [e.g., Behavior Rating Inventory of Executive Functioning (BRIEF), Gioia et al., 2000); Barkley Deficits in Executive Functioning Scale – Children and Adolescents (BDEFS – CA), Barkley, 2012a] (refer to *Appendix C – Commonly Cited Measure of Executive Functioning and Social Emotional Learning Skills*); EF tests (e.g., Children's Gambling Task, refer to *Table A1*); and using the hot/cold analogy within the componential framework (refer to section *Organizational Frameworks of EF*). Second, *inhibition* is a component of both EF and effortful control (Diamond, 2013; Zhou et al., 2012), which can be distinctly measured when EF is measured via EF Tests (see below), but similarly measured when using rating scales to assess EF. For example, the inhibitory control dimension of the Childhood Behavior Questionnaire (CBQ, Rothbart et al., 2001) has items concerning disrupting responses/resisting interference and stopping behavior that are similar to items from the Barkley Deficits in Executive Functioning Scale – Children and Adolescents (BDEFS – CA, Barkley, 2012b) and the Behavior Rating Inventory of Executive Functioning (BRIEF – 2, Gioia et al., 1997 – 2015).

Third, recall EF's *shifting* component is "freely moving from one situation, activity, or aspect of a problem to another as the situation demands" (Gioia et al., 2000, p.321; refer to section *Major Components of Executive Functioning – Shifting*). Thus, EF's shifting is similar to effortful control's attention shifting defined above as re-directing attention (Eisenberg & Zhou, 2016; also see Rothbart & Bates, 2006). EF's BRIEF/ BRIEF-2 (Gioia et al., 1997 – 2015) Shift subscale, which contains items that could be classified primarily as "adjusting to change" and "getting stuck" (see *Table A2*), contains items that are quite similar to items on the attention-focusing and attention shifting subscales of the Childhood Behavior Rating Questionnaire's (CBQ, Rothbart et al., 2001).

Distinctions

One major difference between these constructs concerns the direction researchers have taken with the effortful control literature focusing on socioemotional competence/skills, with a growing interest in academic achievement. By contrast, cognitive and academic development has been a prominent focus in the EF literature, without focusing on social emotional competencies (Zhou et al., 2012). Furthermore, EF is predominantly measured by tests and sometimes rating scales, whereas effortful control is predominantly measured via rating scales and observations (Diamond, 2013).

A second major distinction among these concepts concerns their conceptual foundation concerning the components/subdimensions of *working memory, inhibition (EF tests), shifting (EF tests), and planning/problem solving*. The measurement of EF commonly includes the component of *working memo*ry (refer to section *Executive Functioning as a Predictor of Academic Achievement*). However, items pertaining to working memory are scarce to non-existant in measures of effortful control (Eisenberg & Zhou, 2016; Zhou et al., 2012). In brief, inhibition (EF tests), shifting (EF tests), and planning/problem solving, which share similarities in their definitions among the constructs, by measurement concerns directing the self towards task-oriented problem solving behavior for EF, but regulating oneself more broadly around emotional control for the constructs of self-regulation and effortful control.

Specifically, with respect to *inhibition*, effortful control measures the inhibition or regulation of actions, emotions, or attention (Eisenberg et al., 2004; Eisenberg et al., 2011; Rothbart, 2011) to deal with relevant information (e.g., shift attention, focus attention) to engage in or inhibit behaviors for the purpose of being socially appropriate (Rothbart & Bates, 2006; Rothbart & Rueda, 2005). In contrast, the way inhibition is measured in EF Tests concern lab-based measurement of the execution of and end result of attempts to inhibit or disrupt responses to evaluate more objective (less emotionally based) goal-oriented, problem-solving behaviors. Similarly EF tests that measure *shifting* is distinct from the way shifting is measured via rating scales of either effortful control or

EF as these tests assess speed and ability to understand rules for the purpose of a goal-directed problem solving behavior (refer to *Table A1*).

Last, EF's *planning/problem solving* component differs conceptually from effortful control's executive attention (e.g., correct/detect errors and planning). According to Eisenberg et al.'s (2011) interpretation of the work of Compas, Connor-Smith, and Saltzman (2001), effortful control's executive attention functions to cope with (social) stress, yet EF's planning/problem solving refers to goal-directed problem solving behavior (refer to section *Components of Executive Functioning – Planning/problem solving*). I would further add that EF's planning/ problem solving is also commonly measured via two of the most commonly cited rating scales [i.e., Behavior Rating Inventory of Executive Functioning (BRIEF), Gioia et al., 2000; Childhood Executive Functioning Inventory (CHEXI); Thorell & Nyberg, 2008] and a scale by an esteemed researcher in the field of EF [i.e., Barkley Deficits in Executive Functioning Scale – Children and Adolescents (BDEFS – CA), Barkley, 2012b]. However, the most commonly cited rating scale of effortful control, the Child Behavior Questionnaire (CBQ; Rothbart et al., 1994, Rothbart et al, 2001), does not contain a planning/problem solving component.

Case Examples of Conceptual Overlap

Zhou, Chen, and Main (2012) initially called for an integrated model of self-regulation by combining EF and effortful control into the same construct. However, this position was later reconsidered by Eisenberg and Zhou (2016) who call for the literature to address this overlap due to conceptual and measurement issues that arise when these constructs are not properly distinguished. I will continue to elaborate on this discussion by providing two patterns I noticed in the EF literature with respect to models with conceptual overlap that apply to the area of (a) self-regulation and effortful control and (b) the concept/construct of learning related skills.

On Self-Regulation and Effortful Control. There is an underlying assumption in the SEL literature that SEL Skills improve academic performance through the acquisition of cognitive skills (Goleman, 1995; Elias et al., 1997; Hoffman, 2009). Thus, with regards to the first pattern that emerged in a review of the literature considering the relationship between EF and SEL Skills, some models suggest cognitive skills, including EF, act as a mediator of Social (i.e., SEL) Skills (Yeates et al., 2007; Yates, 2013). Such models used Social Information Processing framework (Crick & Dodge, 1990) as a construct (Dodge et al., 2002) that integrates cognitive-executive, affective, and social problem solving (Moran et al., 2015; Wolfe et al., 2015). For example, using the hot/cold EF paradigm, Yeates et al. (2007) proposed a model whereby EF is a mediator of various developmental states and traits, including social problem solving skills. Beauchamp and Anderson (2010) extended Yeates et al.'s (2007) model by suggesting cognitive structures (e.g., neuropathways) are involved in the development of social skills and include additional global variables such as the environment, temperament, and genetics.

Validity issues can arise in the application of these models, particularly in cases where EF and effortful control are used as interchangeable constructs. In a test of these models in children with various levels of traumatic brain injury, Ganesalingam et al. (2011) measured EF with three EF tests and two rating scales, the BRIEF (Gioia et al., 2000), and the *effortful control* subscale of the Childhood Behavior Questionnaire (CBQ; Putnam & Rothbart, 2007). Anderson et al. (2013) used the same sample as

Ganesalingam et al. (2011), but with the inclusion of a control group. Anderson et al.'s (2013) findings suggest cognition, including EF as measured in the study predicts social development (interpersonal and social competence, social interaction, and social participation). The problem with this research concerns construct validity: both authors included the effortful control subscale of the CBQ to measure EF. As previously demonstrated earlier in this section, EF is not effortful control. This is not to say that EF and the SEL constructs of self-regulation or effortful control cannot be measured in the same model; some researchers do this successfully, minimizing validity issues (e.g., Blair & Ursachi, 2006; 2011; Hoffman, Schmeichel, and Baddeley, 2012). Rather, the take-away of this subsection is that the application of these constructs in the same model must apply a careful awareness of how these constructs can potentially overlap, and avoidance of such overlap at all cost.

On Learning-Related Skills. I would suggest herein that conceptual overlap issues between EF and SEL Skills is present in another body of literature, that of the "learning-related (social) skills" (e.g., see McClelland & Morrison, 2003). Learning related skills is commonly referenced in this body of literature based on the work of McClelland, Acock, and Morrison (2006), who define this as a single construct that measures diverse aspects of multiple constructs: behavioral self-regulation, social competencies (e.g. responsibility, independence, and cooperation), and "behavioral and social manifestation of EF skills", such as listening to and following directions, and organizing materials. A corresponding measure called Head-Toes-Knees-Shoulders has been developed (HTKS; Cameron-Ponitz et al., 2008) and used by some researchers (e.g., Cadima et al., 2016). Learning-related skills, sometimes referred to as learning related behaviors, is a concept that is measured in a variety of ways in the literature (e.g., see Ansari & Gershoff, 2015; Cerda et al., 2014; Sasser et al., 2015; Stipek et al., 2010). It is the way that learning related skills/behavior is measured when included in models that contain EF that can be problematic.

For example, in some of the learning related skills literature, EF and self-regulation are sometimes interchangeable discussed as if they were the same construct (e.g., McClelland et al., 2007; McClelland & Cameron, 2012). However, as previously discussed earlier in this section, while EF and self-regulation share select similarities, they are distinct constructs. While there is some acknowledgement that EF may underlie learning-related behaviors/skills (Stipek et al., 2010), when EF and learning related skills are measured in the same model, I would suggest there is a problem of conceptual overlap (e.g., Sasser et al., 2015). Similar to other definitions of EF in the literature (refer to section *Definitions of Executive Functioning*), Cameron-Ponitz, McClelland et al., (2009) defined EF as: "the manifestation of executive function skills in overt, observable responses in the form of children's gross motor actions…(which) involves multiple components of executive function including attentional focusing, working memory, and inhibitory control" (p. 605). However, the construct of learning related social skills as defined above by McClelland et al. (2006) includes a description akin to two of the most commonly cited components of EF (Baggetta & Alexander, 2016): working memory ("listening to and following directions") and planning ("organizing materials"). These components are also represented in two of the most commonly cited childhood rating scales of EF (Catale et al., 2015): the Behavior Rating Inventory of Executive Functioning (BRIEF; Gioia, et al., 2000) and the Childhood Executive Functioning

Inventory (CHEXI; Thorell & Nyberg, 2008) (refer to *Appendix C – Commonly Cited Measure of Executive Functioning and Social Emotional Learning Skills*). Thus, by including components of EF (i.e., working memory and planning) in the construct of learning related social skills, while simultaneously measuring EF (e.g., working memory, inhibition) in the same model, conceptual overlap can occur, leading to validity issues. Additionally, program effects (i.e., academic achievement) may be misattributed to SEL Skills when in fact they are due to EF.

Construct Overlap: Ratings Scales Item Similarities

Herein I propose that the lack of construct clarity in the literature is reflected in overlap at the level of rating scales with items from the most commonly cited SEL Skills rating scales overlapping with items on EF rating scales, in effect blurring constructs.

The analysis I undertook included some of the most commonly cited childhood rating scales of EF (Catale et al., 2015) and SEL Skills, which are described in detail in Appendix C (*Commonly Cited Measure of Executive Functioning and Social Emotional Learning Skills*). Specifically, I examined all items from the teacher and parent versions of the following EF rating scales: (a) Behavior Rating Inventory of Executive Functioning 2 BRIEF – 2 ; Gioia et al., 2015); (b) Childhood Executive Functioning Inventory (CHEXI; Thorell & Nyberg, 2008); and (c) the Barkley Deficits in Executive Functioning Scale – Children and Adolescents (BDEFS – CA; Barkley, 2012b). I concurrently examined all items from the teacher-versions of the following SEL Skills rating scales: (a) Devereux Student Strengths Assessment – Second Step® Edition (DESSA – SSE; Devereux Center for Resilient Children, 2012); (b) Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997); (c) the Social Skills Improvement

System Rating Scales (SSIS – RS; Gresham & Elliott, 2008); and (d) the Social Skills Improvement System SEL Edition Rating Form (SSIS–SEL RF;Gresham & Elliott,2017).

A qualitative analysis was conducted identifying similarly worded items of the above noted SEL Skills rating scales to EF rating scales in relation to my understanding of the definition of each component of EF provided earlier (refer to section *Major Components of Executive Functioning*). Table A2 lists select items from said EF rating scales with select items from said SEL Skills rating scales as I perceived them to overlap. Based on the way components have been defined or measured in the literature, I have assigned categories or "sub-definitions" to each component as a means to clarify this comparison across hundreds of items among the EF and SEL rating scales. For example, working memory is broken down into the "sub-definitions" of remember, steps of a task, and learn from experience. Note, however, that the replication of rating scale items was limited due to copyright permission issues (as outlined in *Appendix D*).

I have also quantified the similarities among the rating scales. There was a varying amount of overlap among the items on these scales. This overlap was quantified by noting the percentage that these SEL Skills scales overlapped with the aforementioned EF scales. The percentage of this overlap is as follows: 36% (13 out of 33 items) on the DESSA-SSE; 16% (4 out of 25 items) on the SDQ; 5% (4 out of 76 items) on the SSIS – RS; and 3% (2 out of 76 items) on the SSIS – SEL RS. Thus, there is wide variation in the extent to which SEL rating scales overlap with components of EF as measured by EF rating scales, ranging from minimal, and therefore inconsequential (i.e., SSIS) to substantial. Clearly, understanding the specific and potentially synergistic manner of SEL and EF domains requires elucidation at the measurement level.

CURRENT STUDY

Study Aims

Two main research questions (or Aims) will be addressed in the current study for the purpose of filling previously noted gaps in the literature concerning the relationship among EF, Prosocial Skills (as a proxy for SEL Skills), and academic achievement. In short, the purpose of Aim 1 is to address the gap in the literature of the bidirectional relationship between EF and Prosocial Skills. The purpose of Aim 2 is to address the gap in the literature of the longitudinal relationship between EF, Prosocial Skills, and academic achievement, where Reading and Math are two proxies used for outcome variables representing academic achievement. The ultimate goal of answering these two Aims is to contribute to the prevention science literature (and in turn, translational science) in unpacking these two inter-related constructs (Masten & Cicchetti, 2010).

Both Aims are addressed using a developmental cascade model (refer to Figure B1), which is a broad term used to examine the spillover effect that may occur from one competence, adaptive skill or behavior, to another adaptive skill or behavior at a later period of time. Change in one area of functioning is viewed as resulting in change in another area in a later developmental period (Sameroff, 2000) via progressive associations. This spillover could be conceptualized in a bidirectional way to assess whether SEL is an adaptive skill promoting EF and whether EF is an adaptive skill promoting SEL. Put another way, periods of time are essentially measuring units of change. That is, at any given moment, a snapshot could be taken (or a survey administered) to assess children’s development in any given skill, such as EF or Prosocial Skills. However, it is only through examining the bidirectional relationship between two skills that we are able to tease out whether any given skill is developing on its own, or in conjunction with another, closely related skills. Thus, this developmental cascade model will test whether these adaptive skills have an effect, or spillover, on each other (research question 1), as well as the spillover between EF and Prosocial Skills, respectively, on academic achievement over time (research question 2).

Study Strengths

The current study has two main strengths that bolster my ability to address the previously identified gaps in the literature. First, I use a developmental cascade model across four waves of data to examine said gaps in the literature. More specifically, a series of longitudinal auto-regressive cross-lagged panel models are proposed to examine both research questions. This type of model "examines the predictive (direct regression) relations among the latent constructs over time" (Little, 2013). One advantage of a cross-lagged panel model is that "arguments based on theory, prior research, and the statistical control of plausible confounds to the validity of a panel model lend strong implications for drawing a valid causal conclusion" (Little, 2013, p.181). Second, the results are generalizable to a broad population of North American school children because the sample is large and contains non-clinical, ethnically diverse American school children.

Hypotheses

Five hypotheses are explored in the current study, subdivided by study aim.

Aim 1: Bidirectional Relationship

Hypothesis 1 (H1). EF and Prosocial Skills will positively and reciprocally influence each other across time points, controlling for the stability of each construct respectively.

Hypothesis 2 (H2). Given that both EF and Prosocial Skills increase as children age and that these skills are expected to be moderately correlated constructs, it is hypothesized that the magnitude of the reciprocal association between EF and Prosocial Skills will strengthen over time.

Hypothesis 3 (H3). Children's gender will influence EF and Prosocial Skills.

Aim 2: Academic Achievement

Hypothesis 4 (H4). EF will directly and positively influence academic achievement.

Hypothesis 5 (H5). Prosocial Skills will directly and positively influence academic achievement, but EF will be a stronger predictor of academic achievement than Prosocial Skills.

METHOD

Sample and Demographics

Data used were limited to students in the control group, who participated in a larger randomized controlled intervention study (*N* = 7,200 students) that assessed the effects Second Step® (Low et al., 2017; 2018). The sample for the current study is derived from the control group. Low et al. (2015) reported no significant differences between treatment and control groups on baseline measures.

The resulting student sample consisted of 3,029 students in Year 1 and 3,078 students in Year 2. Two cohorts of students were followed in each year, such that at Time 1 students were either in Kindergarten, referred to as Cohort 1 (*n* = 1,435), or Grade 1, referred to as Cohort 2 (*n* = 1,594). Additional information is provided in the following section (refer to section *Recruitment and Retention*). Student demographic data (e.g., age, grade, gender, ethnicity, English language learner, lunch program distribution, geographic location) were collected from teachers (refer to *Table A3*). With regards to geographic representation the majority (67%) of the students were located in school districts in Washington (i.e., Renton, Federal Way, Lake Washington, Mukilteo, and North Kitsap), with the remainder of the students from the district of Mesa in Arizona. The gender and ethnicity distribution of school-age children in the United States was well represented in the sample (see Low et al., 2015). Data on the percentage of students in each school who were eligible for free and reduced lunch programs was collected as a means to assess socioeconomic status. Sixty-seven percent of students in the sample were located in schools where 50% of more of the children were eligible for free and reduced lunch programs. School's lunch program status is on par with statistics that suggests roughly half of school children in the United States are eligible for free and reduced lunch programs (US Department of Education, Table 204.10, 2018). The percentage of students eligible for free and reduced lunch programs in each school is similar to the percentile reported in census data on elementary schools.

Teachers self-reported demographic data (e.g., age, gender, ethnicity, education level, grades taught), which was collected at Time 1 of the study (refer to *Table A4*). The study reported on 151 teachers across both study sites ($n_{AZ} = 48$; $n_{WA} = 103$). Data suggests gender, ethnicity, and education were relatively well representative of teachers in the US (McFarland et al., 2019).

Procedures and Design

Data Collection

Data used in the present study includes data that were collected at four assessment waves over two years. Data for Year 1 were collected between October 10 and November 6, 2012 across the various school districts (Time 1) and again between April 22 and May 31, 2013 (Time 2). Data for Year 2 were collected between September 30 and November 8, 2013 (Time 3) and again between April 14 and May 30, 2014 (Time 4).

In accordance with approval by the institutional review boards at Arizona State University and the University of Washington, as well as district policies, passive consent was obtained for students and teachers in the spring of 2012 from schools in urban and rural areas across two states, Washington and Arizona, over six school districts.

Recruitment and Retention

In terms of enrollment, of the seven districts that were approached, six districts participated ($n_{AZ} = 1$; $n_{WA} = 5$). Across these six districts, sixty-one schools ($n_{AZ} = 20$;

n_{WA} = 41) participated. Classrooms from each school were also randomly selected with an average of five and six classrooms per school for Arizona and Washington, respectively. The schools were randomly assigned to either the treatment (*n* = 30) or the control condition (*n* = 31). Schools in the control condition received the intervention post study. As previously noted, this study only uses data for the control condition. Additionally, while the larger study collected data on students who were enrolled in Kindergarten, Grade 1, and Grade 2 at Time 1, with these students followed longitudinally, due to the small sample size of students in Grade 2 at Time 1, these students were excluded from the analyses for the purposes of this study.

Passive consent was used to recruit students. Data were collected from 4,450 students across one or more time points with only 310 students refusing consent. Another 500 students were deleted from the existing database because they only had data present at either the first or fourth time point, which could not be used to predict change. Thus, a total of 3,640 students had data across one or more time points, with 1,981 of these students having data across all four time points. In terms of attrition, 848 students (20.5%) exited the study after Year 1 (n_{AZ} = 343; n_{WA} = 505), while 744 students (18%) entered the study in Year 2. Thus, student attrition rates in Year 1 of the study were largely compensated by an influx of additional students in Year 2 of the study. These attrition rates are on par with other school-based studies and are in part attributed to residential moves leading to a change in schools (Vuchinich et al., 2012). Student attrition rates in this study were similar to attrition rates reported in previously published research which is drawn from the lager database (intervention and control students) from which the database for the current study is derived (see Low et al., 2019).

Compensation

Schools and staff were compensated as follows: school liaisons received $250 per year to distribute materials, coordinate data collection times, track implementation, and communicate with teachers; and teachers received $5 per online survey per student with a $25 bonus if surveys for their class were completed within three-weeks.

Measures

Executive Functioning (EF)

The current study's database did not contain a measure of EF. However, in the aforementioned qualitative analysis that was conducted comparing widely cited SEL skills and EF rating scales (refer to *Table A2*), there was overlap among some of the items; some SEL skills items closely resembled the wording and/or meaning of an item on a given EF rating scale across the aforementioned five most commonly measured EF components (i.e., working memory, inhibition, shifting, planning/problem solving, and emotional regulation). Initially, 19 items in total were selected across the two SEL scales from the current study's database: the Devereux Student Strength Assessment (DESSA – SSE; Devereux Center for Resilient Children, 2012) and the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997). Of the 19 items, 14 items were derived from the teacher version of the DESSA – SSE (Devereux Center for Resilient Children, 2012), a five-point scale with responses ranging from "never" (0) to "very frequently" (4). The five remaining items were derived from the teacher version of the SDQ (Goodman, 1997), a three-point scale with response options of "not true" (0), "somewhat true" (1), and "certainly true" (2) (refer to section *Measures – Prosocial Skills*). Four of these five items, which were negatively worded, were reverse coded to reflect the positively worded items from the DESSA – SSE (Devereaux Center for Resilient Children, 2012). For an overview of these 19 items, refer to items with the citation of the DESSA – SSE (Devereaux Center for Resilient Children, 2012) and the SDQ (Goodman, 1997) in the right column of Table 2. The DESSA – SSE and SDQ rating scales are judged to have good psychometric properties as described in Appendix C (*Commonly Cited Measure of Executive Functioning and Social Emotional Learning Skills*).

After a process of determining the factor structure of the current study's EF measure (refer to section *Results*), results suggested only 13 of the hypothesized items from the two aforementioned rating scales were represented by two factors. These factors are consistent with the EF childhood rating scale literature that measures items related to working memory, planning, and problem solving on one factor (EF 1), and items related to shifting and emotional regulation on a separate (EF 2) factor (refer to Appendix C – *Commonly Cited Measure of Executive Functioning and Social Emotional Learning Skills*). The items representing these two factors are listed in Table A5. Note that two of the items on EF 2, which are from the Emotional Symptoms and Conduct Subscales of the SDQ (Goodman, 1997), were reverse coded (contrary to the original scale's directions) to correspond with the other items on the EF 2 from the DESSA – SSE (Devereux Center for Resilient Children, 2012) that were positively worded. As such, for both EF 1 and EF 2, the higher the score, the greater the perceived assessed skill level. These same items were administered to students in each grade across all four waves over both years of the study. As shown in Table A6, the reliability of each measure, as represented by Cronbach's alpha is .95 for EF 1 and ranges from .71 to .78 for EF 2.

Prosocial Skills

After a process of determining the factor structure of the current study's Prosocial Skills measure (refer to section *Results*), results suggested four of the five hypothesized items from the prosocial subscale of the teacher version of the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) were representative of prosocial behavior. The SDQ is a behavior rating scale that is comprised of 25-items rated by teachers or parents (two versions) to assess the perceived strengths and difficulties of children aged 4 – 16 years. The items selected for the current study, which are representative of prosociality as a multi-dimensional factor (e.g., sharing, helping, social awareness, relationship skills), are listed in Table A5. Items are scored on a 3-point scale with 0 indicating "not true", 1 indicating "somewhat true", and 2 indicating "true", such that a *higher* score indicates more prosocial behavior. The same version of the scale was administered to students in each grade across all four waves (over two years) of the study.

The SDQ is a widely used measure for screening and research purposes that has been translated into over 60 languages (Stone et al., 2010). The prosocial subscale of the SDQ teacher version was borrowed from Weir and Duveen (1981) Prosocial Behavior Questionnaire, which was judged to have good validity and high internal reliability (α = .93 -.94 across three samples); split half reliability of α = .82 – .85); excellent test-retest reliability ($r = .91$, $p < .001$); and fair inter-rater reliability of ($r = .58$, $p < .01$). Validity of the SDQ (Goodman, 1997) is suggested by high predictive and concurrent validity (Goodman, 1997), as well as criterion and discriminant validity (Goodman & Scott, 1999). According to Stone et al.'s (2010) meta-analysis of construct validity, when examining factor loadings on all five of the prosocial scale items for the teacher version most studies reported factor loadings of .70 or above (*N* = 19,105, *k* = 7). Additionally, Stone et al. (2010) concluded the teacher version's psychometric properties in samples of children aged 4 to 12 (*N* = 131,223) suggests the internal consistency weighted mean (*k* = 26) on the teacher version is good (α =.82) and that the test-re-test weighed mean (*k* = 8) is acceptable (*r* = .79).

As indicated in Table A5 the current study suggests the Prosocial Skills factor has good reliability (α =.77 – .82).

Academic Achievement

Aimsweb curriculum-based measures of oral reading fluency (RCBM) and math calculation (M-CBM) were collected multiple times per year with the fall and spring academic achievement scores used for the purpose of this study across the two years. Graduate research assistants followed standardized administration directions from the aimswebPlus website (www.aimsweb.com).

R-CBM was a grade appropriate short (1 minute) reading passage that was administered individually to each student. The reading passage was scored as correct words read per minute, with the range of words contained in the reading passages increasing in number with each successive grade (refer to *Table A6*). Oral reading fluency (R-CBM) probes are standardized, general outcome measures of reading performance and are highly sensitive to instruction (Fuchs & Fuchs, 1999).

The math test (M-CBM) was group administered and scored individually based on the percentage of math problems correct in the time allotted (8 minutes). Each grade was administered math problems appropriate for their grade level. Math computation probes, such as the M-CBM, have been shown to be a reliable and valid general outcome measure of computation (Thurber et al., 2002).

Student and School Demographics

Teachers provided the demographic data for their students (see Table A3) and self-reported their own demographic data (see Table A4). The remaining demographic data, such as the percentage of students receiving free or reduced-price lunch (see Table A3), were collected using publicly available on-line sources (e.g., NCES website, school district websites).

Covariates

Because a child's age and gender can affect the development of EF and Prosocial Skills in childhood, the current study incorporates both of these factors, with age accounted for by the longitudinal cross-sectional design and gender accounted for as a predictor variable, as per Hypothesis 3 (refer to section *Current Study*).

In the current study, age was measured by subtracting children's date of birth as reported by their teacher at Time 1 from the starting date and year on which data were collected across each of the four waves. Gender was assessed at Time 1 using teacher reports (0 = *male*, 1 = *female*), and is thus a measure of perceived gender.

Analytic Strategy

Preliminary Analyses

Descriptive Statistics. Because the estimator (i.e., WLSMV) for the structural equation analyses described below does not assume multivariate normality, descriptive statistics were performed accordingly. Specifically, preliminary analyses were conducted using SPSS v.26 (IBM Corp., 2019) to identify univariate outliers at the indicator and variable levels and to generate descriptive statistics and frequencies for students by cohort and teachers. Additional descriptive statistics were conducted by cohort for EF 1, EF 2, and Prosocial skill and for Math and Reading. Within and across time correlations among EF 1, EF 2, and Prosocial Skills, including Math and Reading were conducted, which served to assess for extreme collinearity (correlations above .90).

Exploratory Factor Analyses. The number of factors that comprise EF (delineated as EF 1 and EF 2, as previously indicated) and Prosocial Skills were assessed using EFA because when there is uncertainty regarding the number of factors that comprise a construct, an EFA should be conducted (Pett et al., 2003) to reduce the likelihood of needing to respecify the hypothesized model based on results of Confirmatory Factor Analysis (Nesselrodes & Baltes, 1984). EF and Prosocial Skills were estimated by conducting an EFA within the CFA framework (E/CFA) using Mplus v.8.1.5 (Muthén & Muthén, 1998-2017) to develop a more realistic measurement structures prior to conducting a CFA (Brown, 2006). As a means to test for factor distinction between EF and Prosocial Skills and in turn multicollinearity in the structural model, a combined EFA was conducted with all 24 measured variables (i.e., 19 items representing EF and 5 items representing Prosocial Skills).

Rather than conducting an EFA on each of the four waves of data, one EFA was conducted by randomly selecting one wave for each student. Specifically, data was converted from wide to long format in SAS based on variables that determined students' grade. An observation was randomly chosen for each student based on a random sample generator in SAS/STAT® software v.9.4 to select one of the four possible waves of data available per student. Last, the data was split in two such that one half of the sample was randomly selected to create a database for the EFA and the other half of the sample was used for the CFA (refer to section *Measurement Model* below).

Due to the categorical nature of the measured variables that use two different scales (refer to section *Method*), the estimator used was the weighted least square parameter estimates using a diagonal weight matrix with standard errors and mean- and variance- adjusted (WLSMV; Muthen et al., 1997) with Theta parameterization (see Muthén & Muthén, 1998-2017). The WLSMV estimator in Mplus (Muthén & Muthén, 1998-2017) provides more accurate and less biased factor loadings (e.g., Li, 2016). The variance of each factor loading was freely estimated using the Mplus default method by setting the factor loading of the first measured variable to 1.0 and the latent mean to 0.

Measurement Model

To test the measurement model, Confirmatory Factor Analyses (CFAs) were conducted followed by longitudinal CFAs.

Confirmatory Factor Analyses. The factor structure that materialized from the EFA to investigate the factor structure of EF and Prosocial Skills described above was subsequently cross validated on the second half of the sample using confirmatory factor analysis (CFA) in Mplus v.8.1.5 (Muthén & Muthén, 1998-2017). The process used to extract the sample was described above (refer to section *Exploratory Factor Analysis*). Evaluation of model fit was assessed using the following fit statistics: Chi Squared Model Fit (χ^2), Bentler's Comparative Fit Index (CFI) (Bentler & Bonett, 1980), the Root Mean Square Error of Approximation (RMSEA), and the Weighted Root Mean Square Residual (WRMR) values. The WRMR is an alternative fit statistic to the Standardized Root Mean Square Residual (SRMR) that was developed by Muthén and is provided in the Mplus output when using the WLSMV estimator when data contains ordered categorical variables (see Muthén & Muthén, 1998-2017). Values suggestive of a well-fitting model are as follows: a *p*-value greater than 0.05 for χ^2; 0.95 or above for the CFI; and a value lower than 0.90 for the WRMR (Muthén & Muthén, 1998-2017; DiStephano et al., 2018; Yu & Muthén, 2002). RMSEA has a range from mediocre fit (>.08 - < .10) (MacCallum et al., 1996), fair fit (>.05 - <.08) (Browne & Cudeck, 1993), to close fit (< = 0.5) (Browne & Cudeck, 1993).

Longitudinal Confirmatory Factor Analyses. After having conducted a CFA for the hypothesized EF (EF 1 and EF 2) and Prosocial Skills scales, the next analytic step was to test factorial measurement invariance, that is to conduct longitudinal CFAs, using Mplus v.8.1.5 (Muthén & Muthén, 1998-2017) with the original (unsplit) database on the first and fourth assessment waves (for each student). The same model fit statistics (i.e., χ^2, CFI, RMSEA, WRMR) and standards of model fit used previously were applied. The estimator WLSMV was selected because, as previously indicated, the latent variables are categorical, and theta parameterization was used. We used the DIFFTEST in v.8.1.5 (Muthén & Muthén, 1998-2017) to compare nested models as well as change in CFI (> .01) to assess measurement invariance.

Examining longitudinal measurement invariance using the WLSMV estimator is different than examining it using the maximum likelihood estimator. As such, we followed recommendation from Liu and colleagues (2017) to examine longitudinal measurement invariance for each of our three latent variables. The following models were used to test for factorial measurement invariance: (a) configural model to assess whether the factor structure remains consistent across time by fully unconstraining all estimated parameters (configural Model 1) and respecifying the model if needed by partially unconstraining estimated parameters (Model 1b); (b) loading factorial invariance model to assess whether each latent variable was comparable across time by fully constraining loadings to be equal to one another over time (Model 2a), and respecifying the model if needed to partially constrain loadings (Model 2b); (c) threshold factorial invariance by fully constraining the thresholds (Model 3a), and respecifying the model if needed to partially constrain the thresholds (Model 3b); and (d) unique factor invariance by constraining all unique variances to 1.0. The unique factor invariance model is similar to the strict invariance model for continuous indicators.

Structural Model

After establishing full or partial measurement invariance among the three latent variables, a full longitudinal CFA model was fit with each latent variable over time. The Auxiliary command in Mplus v.8.1.5 (Muthén & Muthén, 1998-2017) was used to include manifest variables of students' gender, Math, and Reading. The full model was run and the factor scores were outputted and saved for each of the three variables. SAS/STAT® software v.9.4 was used to change the measurement of time by converting the data from wave to semester grade. Specifically, this accelerated longitudinal design is a planned missing design that leveraged the overlap between the two age cohorts and reorganized the data for the purpose of examining differences across longer periods of time and across semesters, which is more developmentally appropriate than wave. This method combined the two cohorts across four waves into one single cohort across six waves to examine a longer period of time (from the fall of Kindergarten through the spring of Grade 2). Once the data was restructured, the hypothesized longitudinal structural models were fit using Mplus v.8.1.5 (Muthén & Muthén, 1998-2017). Specifically, one model was fit for math achievement (Model 1 – Math) and another model was fit for reading achievement (Model 2 – Reading). All estimates were reported using the completely standardized estimates.

RESULTS

The missing data and descriptive statistics are presented for the study's main variables. Of note, while EF was conceptualized as one latent variable, in keeping with the results of the factor analyses presented later in this section and consistent with the literature reviewed of the most commonly cited rating scales of EF (refer to *Appendix C Commonly Cited Measure of Executive Functioning and Social Emotional Learning Skills*), this study examines EF as multifactorial. As such the missing data and descriptive statistics are presented for EF 1, which represents EF's components of working memory and planning/problem solving, and EF 2, which represents EF's components of shifting and emotional regulation (refer to section *Major Components of Executive Functioning*).

Missing Data

Missing data is presented below with respect to the sample's descriptives as well as dependent variables. The pattern of missing data is also discussed.

Sample

With regards to the study's variables (EF 1, EF 2, Prosocial Skills, Reading, and Math), missing data ranged from 17.8% to 31.6% for Cohort 1 and 13.4% to 30.2% for Cohort 2 across the four waves (see Tables A6 and A7). This level of missing data is consistent with missing data reported using not only the control data (as in this case), but intervention data in the original database by Low et al. (2019). Student demographic variables indicated missing data ranged from 0% to 29.5% for Cohort 1 and from 0% to 20.4% for Cohort 2 (see *Table A3*). Teacher demographic variables (collected for the whole sample) indicated that seldom (0% - 6%) was data missing (see *Table A4*).

The marginal missing teacher demographic data was suggestive of a good adherence to data collection and data entry, while the high amount of student data was attributed to student absence during the day of the assessment (e.g., illness, appointments) or mobility. That is, high student mobility is common in schools where there is a large percentage of students participating in free and reduced lunch programs (Vuchinich, Flay, Aber, & Bickman, 2012), such as in the case with the present study.

Pattern

Missing data patterns from prior research using this database suggests the pattern was missing at random (Cook et al., 2018; Low et al., 2015, 2016, 2019).

Descriptive Statistics

Outliers

Because data screening procedures can affect Type 1 and 2 error (Tabachnick & Fidell, 2012), the variables were examined for univariate and multivariate outliers. An assessment of outliers was made at the item level for EF 1, EF 2, and Prosocial Skills and at the variable level for these variables, in addition to Reading and Math.

Item Level. Before the factor structure of EF and Prosocial Skills was finalized via factor analysis for the purpose of the structural model analyses herein using Mplus v.8.1.5 (Muthén & Muthén, 1998-2017), the items comprising these scales were converted to *z*-scores using SPSS v.26 (IBM Corp., 2019) solely for the purpose of locating extreme scores (z = >+-3.29) one standard deviation below or above the mean. However, due to the sample size (n_z = 1454), scores slightly outside of this range were not considered extreme (Tabachnick & Fidel, 2006). The outliers were within range or marginally outside of range among the items that comprised the EF 1 (z = -2.4 – -3.7), EF 2 (z = -2.7 – -4.3), or Prosocial Skills (z = -1.9 – -3.2) scales.

Variable Level. Dependent variables with potential univariate outliers were considered by converting the variables across the four Time points in the model (EF 1, EF 2, Prosocial Skills, Math and Reading) to z-scores to locate extreme scores one standard deviation below or above the mean. Again, based on the sample sizes (n_z = 1454) the following z scores are not considered extreme for most of the study's variables across the waves (Tabachnick & Fidel, 2006): EF 1 (z = -3.5 – 1.6); EF 2 (z = -4.8 – 0.7); Prosocial Skills (z = -3.4 – 0.9); Math (z = -4.4 – 9.3); and Reading (z = -2.1 – 5.4). The only possible extreme z-score was for Kindergarten Math scores, which ranged from -0.5 to 9.3 at Time 1 (fall), while all the other z-scores for Math across the other time points ranged from -4.4 to 3.5.

Frequencies and Descriptives

The study's main variables were summarized separately for each of the two cohorts. The descriptives for EF 1, EF 2, and Prosocial Skills are provided in Table A6 and for Math and Reading in Table A7. Only the Math and Reading variables were highly kurtotic (30.8 and 21.4, respectively) for Cohort 1 (Time 1, fall). This is not surprising given the variable skills level upon entry into Kindergarten, with extreme differences in children's math and reading abilities declining over time (becoming less kurtotic) as the children move forward in grade level (refer to *Table A7*). However, it should be noted that kurtosis was not that important of a consideration because the WLSMV estimator selected for the structural model to account for the categorical nature of these variables was not restricted to the same set of distributional assumptions of multivariate normality (i.e., skewness and kurtosis, linearity, and homogeneity of variance) as other estimators (e.g., ML).

Correlations

Within and across time correlations were conducted using SPSS v.26 (IBM Corp., 2019) by Cohort (see *Tables A8* and *A9*). Within time correlations also functioned as a way to examine extreme multicollinearity (r >.90), which can affect construct validity. Table A8 presents the within and across time correlations among EF 1, EF 2, and Prosocial Skills. Most correlations were statistically significant at the .01 level and ranged from small to large size effects, no larger than r = .71. Table A9 presents the correlations between each of these variables and Math and Reading. Most correlations were statistically significant and ranged from small to medium size effects but were no larger than r = .45.

Exploratory Factor Analysis

Combined Factors Model

Based on a priori work an exploratory factor analysis (EFA) was conducted in Mplus v.8.1.5 (Muthén & Muthén, 1998-2017) using the first split half sample (n = 1,130) previously described (refer to section *Analytic Strategy*) for the purpose of evaluating construct distinction between EF and Prosocial Skills (refer to *Table A2*) in an effort to minimize the risk of multicollinearity in the structural model. In total, 24 measured variables (i.e., rating scale items) were included in the analysis, based on: (a) 19 measured variables of the hypothesized EF factor as listed in the right most column of Table 2 with citations of the DESSA – SSE (Devereaux Center for Resilient Children, 2012) and SDQ (Goodman, 1997); and (b) five measured variables of the hypothesized

Prosocial Skills factor derived from the prosocial scale factor of the SDQ (Goodman, 1997). These two hypothesized scales were described in the *Measures* section.

The number of factors specified was between one and four. The decision on how many factors to extract was based on a priori theory and research. Specifically, the five measured variables representing Prosocial Skills, extracted from the prosocial subdimension of the Strengths and Difficulties Questionnaire (Goodman, 1990), were hypothesized to load onto one factor. The 19 measured variables representing five components of EF (*refer to section Measures*) were hypothesized to load onto two additional factors (Gioia et al., 2000, 2015; Thorell & Nyberg, 2008) (*refer to Appendix C Commonly Cited Measures of Executive Functioning and Prosocial Skills*). While an EFA with measured variables representing EF (2 factors) and Prosocial Skills (1 factor) was expected to yield a three-factor solution, a four-factor solution was also considered.

Factor Loadings. The results of the one to four factor solutions are interpreted below based on Comrey and Lee's (1992) guidelines for evaluating factor loadings [e.g., very strong (r = .64 or above; 40% shared variance); excellent (r = .71 or above; 50% shared variance)].

1-Factor Solution. Interpretation of the 1-factor solution was inconsistent with prior research and theory that conceptualizes EF as a distinct construct from Prosocial Skills. Of the 24 measured variables, 20 variables had "very good loading" (.65 or greater, 40% shared variance). The remaining four measured variables, which represented the emotional regulation component of EF, loaded less strongly onto the factor than the other measured variables (λ = .37 – .63).

2-Factor Solution. The interpretation of the 2-factor solution was also inconsistent with prior research and theory. The first factor consisted of all the measured variables hypothesized to represent the working memory/plan/problem solving component of EF. These measured variables all had excellent loadings (λ = .65 – .95) onto the first factor with low cross-loadings (λ = -.00 – .34) on the second factor, so as to suggest the factor is meaningfully distinct (Pett et al., 2003). The second factor consisted of all of the measured variables hypothesized to represent the emotional regulation and shifting components of EF and Prosocial Skills. While there was some theoretical rational to believe emotional regulation component of EF and Prosocial Skills may load on the same factor, there was no theoretical reason for shifting to load onto this factor. These measured variables have excellent loadings (λ = .71 – .82) with low to moderate cross-loadings (λ = -.42 – .43) onto the first factor.

The upper range of the cross-loadings for Factor 2 are largely represented by two of the three measured variables hypothesized to represent the inhibition components of EF. While these measured variables (DESSA 21 and SDQ21) loaded most strongly onto the second factor (λ = .52 and .51, respectively), they have cross-loadings onto the first factor (λ = .43 and .37, respectively).

3-Factor Solution. The interpretation of the 3–factor solution was largely consistent with prior research and theory, and therefore the results are described in more detail. Based on the pattern matrix, Table A10 shows loadings for this 3-factor solution with all 24 items.

Measures loading dominantly on Factor 1, called EF 1, which comprised 10 of EF's measured variables hypothesized to represent working memory (all 3 items), inhibition (1 of 3 item), and planning/problem solving (all 6 items) with loadings of these measures on Factor 1 ranging from .51 to 1.1 with some cross-loadings also observed on the other two factors as follows: $\lambda_{\text{Factor 2}}$ = - 0.34 – 0.38; and $\lambda_{\text{Factor 3}}$ = - 0.18 – 0.20. This cross-loading range was being driven up by one item, DESSA 21 (.38), which represents EF's inhibition. It was inconsistent with a priori scale development (e.g., CHEXI and BRIEF) for measured variables representing the inhibition component of EF to load onto Factor 1, as they typically load with Shifting and Emotional Regulation items.

Factor 2, called EF 2, comprised six of EF's measured variables of shifting (both items) and emotional regulation (4 of 5 items). Factor loadings for measured variables loading dominantly on Factor 2 ranged from .60 to .99 with some cross-loadings also observed on the other two factors as follows: $\lambda_{\text{Factor 1}}$ = - .11 – .38; and $\lambda_{\text{Factor 3}}$ = - .23 to .59. Specifically, both measured variables representing the shifting component of EF (DESSA – SSE 31 & 36) had moderate cross-loadings onto Factor 1 (λ = .36 and .39), as did one of the measured variables representing emotional regulation of EF (DESSA – SSE 28). While Factor 1 also represents EF, the components of shifting and emotional regulation in other rating scales typically load onto their own factor (e.g., Behavior Rating Index of Executive Functioning; BRIEF, Gioia et al., 2000).

Factor 3, called Prosocial Skills, comprised all five of Prosocial Skills' measured variables, with excellent loadings (λ = .67 to .78) with no cross-loadings on non-hypothesized factors exceeding 0.25.

To summarize the 3-factor solution the measured variables had the strongest loadings onto their hypothesized factors (refer to *Table A10)* with most loadings deemed very strong (.64 or above; 40% shared variance) or excellent (.71 or above; 50% shared variance) as per Comrey and Lee's (1992) guidelines for evaluating factor loadings. The data suggests there are moderate to large significant (p >.05) correlations among the three factors ($r_{\text{Factor1}\cap 2}$ = .66; $r_{\text{Factor1}\cap 3}$ = .63; $r_{\text{Factor2}\cap 3}$ = .52), but the size of the correlations are not suggestive of multicollinearity (Tabachnick & Fidell, 2007).

4-Factor Solution. In the four-factor solution, only three distinct and meaningful factors emerged. Nearly all of the measured variables loaded onto the same factor as identified in the 3-factor solution with low cross-loadings onto a second (non-hypothesized) factor.

Fit Indices. As shown in Table A11, the chi-squared value did not fit the data well (significant *p*-value) for any of the four models tested (i.e., 1 – 4 factors). However, as the number of factors increased, the value decreased for the RMSEA and WRMR and the value increased for the CFI fit index, suggesting the model with the greatest number (i.e., four) of factors fit the data best. However, as described above, in examining the factor pattern matrix for each model, it was determined that a 3-factor model was the most interpretable model. Overall, the fit indices for the 3-factor model were indicative of a well-fitting model (CFI .974, WRMR .057, RMSEA, .096).

Measurement Model

Prior to conducting analyses to test the structural model, the measurement models developed using EFA were tested in the second split half sample (a fresh sample) using confirmatory factor analyses (CFA) and longitudinal confirmatory factor analyses and longitudinal measurement invariance was evaluated.

Confirmatory Factor Analysis

Based on the results of the EFA described above a CFA was conducted using Mplus v.8.1.5 (Muthén & Muthén, 1998-2017) on the second half of the split half sample (*n* = 1,151) previously described (refer to section *Analytic Strategy*). The measured variables that had cross-loadings of .30 or higher onto their non-hypothesized factor [DESSA – SSE 21, 28; SDQ 8 and 24 (Reverse Coded); and SDQ 21] were dropped. The exception was that the measured variables representing EF's (component of) shifting, which had cross-loadings of .36 and .39 on Factor 1 (the other EF factor) in the EFA, were kept in the final model and allowed to load only on Factor 2 to create a combined shifting/emotional regulation factor, consistent with theory and existing rating scales.

Additionally, because the result of the EFA suggested that one measured variable representing EF's inhibition loaded onto the first factor (labeled working memory/planning/problem solving), which is inconsistent with other childhood rating scales of EF in which inhibition loads with shifting and emotional or behavioral regulation (Factor 2), this measured variable was also dropped from the analysis (the CFAs conducted that forced one or both of the inhibition measured variables onto Factor 2 had an RMSEA and WRMR that fared equally well or worse, respectively).

As previously described (refer to section *Method*), the items representing the measured variables for the EF factors (Factors 1 and 2) are conveyed in Table A5. Factor 1 (referred to as EF 1) comprised of ten items represented by measured variables hypothesized to capture EF's components of working memory (3 items) and planning/problem solving (7 items). Factor 2 (EF 2) comprised of four items represented by measured variables hypothesized to capture EF's components of shifting (2 items) and emotional regulation (2 items). Factor 3 (PS) comprised of all five items representing Prosocial Skills, as previously described.

While the results suggested a statistically significant chi-square ($\chi^2(132) = 1421.85$, $p < 0.001$), the other fit indices taken together are generally indicative of acceptable model fit. The WRMR (2.09) was above 1.0, but the other fit indices were indicative of reasonable fit [CFI = .98; RMSEA = .09 (CI = .09, .10)].

Given the exploratory nature of this paper that aimed to test the structural model by using items from a pre-existing database to create measures representative of EF and Prosocial Skills, I proceeded to test longitudinal measurement invariance for each of these constructs.

Longitudinal Confirmatory Factor Analyses

Each of the three constructs, or factors of EF 1 (Working Memory/Plan/Problem Solving), EF 2 (Shifting/ Emotional Regulation), and Prosocial Skills were assessed for configural, weak (loadings), strong (thresholds), and unique (strict) factorial invariance for the purpose of empirically assessing if each factor was measured consistently across time. The results are presented in Tables A12. As described (refer to section *Analytic Strategy*) factorial invariance was assessed for waves 1 and 4.

Prior to discussing these results, first a word about the way measurement invariance testing was conducted. The determination of whether to conduct successive levels of invariance testing, which may or may not require model respecification, was determined using the DIFFTEST option in Mplus v.8.1.5 (Muthén & Muthén, 1998-2017) because the model was specified using the WLSMV estimator given the ordered categorical nature of the items/indicators. While a non-significant chi-squared value was preferred, as described below, the determination of whether to proceed with assessing factorial invariance of a given model ultimately rested with the other fit indices.

The incremental (i.e., CFI) and comparative (i.e., RMSEA, WRMR) fit statistics can be used to assess model comparison in conjunction with the chi-squared value with the more complex model (i.e., subsequent model) preferred when the fit of the model is improved (Kline, 2011). Of these fit indices the CFI value was the second most important determinant in this process in proceeding with successive levels of invariance testing. The last column in Table A12 indicates the numeric value of the change in CFI between models. If the chi-squared value remained significant following any stage of the respecified model (e.g., configural, weak, strong), then proceeding with factorial measurement invariance was determined by considering a CFI value that was greater than or equal to 0.95, indicative of acceptable model fit, in conjunction with a change in the CFI value that was less than or equal to 0.01 (see Putnick & Bornstein, 2016 for a review). This is in line with Cheung and Rensvold (2002) who argue that this minimal change in the CFI suggests that a set of constrained parameters is the same across time.

Below the results of measurement invariance are reviewed in detail for EF 1, but because the other constructs (EF 2 and Prosocial Skills) had similar results (see *Table A12*), to avoid redundancy, the results reported for these constructs are condensed.

Executive Functioning 1 (EF 1).

Model 1 Configural Invariance. To test whether the theorized factor model of EF 1 with nine measured variables, or indicators (see *Table A5*), fit the data across the two waves (waves 1 and 4) a fully unconstrained configural (baseline) model was first specified (see *Table A12*). The model estimation terminated normally (n = 2,931). As shown in Table A12, the null hypothesis that the model-implied variance/covariance matrix and mean vectors are equal to the population variance/covariance matrix and mean vector was rejected ($\chi^2(132) = 1674.01$, $p < .001$). In other words, based on the chi-squared value I did not retain the model because the differences between the population and model-implied variance/covariance matrices and mean vectors are not minimal. However, in examining the other fit indices there was evidence of a well-fitting (i.e., CFI = .989; WRMR .029) and fair fitting (i.e., RMSEA .063) model. As a result of examining these other fit indices, I proceeded with invariance testing.

Model 2 Weak Invariance. Weak factorial invariance was assessed by running this model with fully constrained loadings of the indicators across the two waves. The model estimation terminated normally. The DIFFTEST option was used to determine if the weak invariance model had significantly worse fit than the configural model (Wu & Estabrook, 2016). The value was statistically significant ($\chi^2(8) = 28.55$, $p < .001$), suggesting that the model with constrained loadings had significantly worse fit than the configural model. Thus, the null hypothesis that the models fit equally well was rejected. However, the other fit indices indicated that the weak invariance model did not differ from the configural model given there was no change in the CFI (see *Table A12*). As such, Model 2 passed the test of weak factorial invariance.

Model 3 Strong Invariance. In running this model with fully constrained means and intercepts of the indicators to test for strong factorial invariance, the aforementioned model constraints (of the loadings) were included. The model estimation terminated normally. Again, the DIFFEST option was used to determine if the strong invariance model had a significantly worse fit than Model 2 – the weak invariance model (Wu &

Estabrook, 2016). The value is significant, suggesting the model with fully constrained means and intercepts is not a good fit to the data and had significantly worse fit than the weak invariance model ($\chi^2(26) = 115.10$, $p < .001$). Thus, based on the chi-squared the null hypothesis that the models fit equally well was rejected. On the other hand, the other fit indices indicated the model is a good fit to the data with a minimal change (.001) in the CFI (see *Table A12*).

Model 4 Unique Invariance. In running this model with fully constrained residual variances to test for unique factorial invariance, the aforementioned model constraints were included (i.e., loadings and thresholds). Again, the DIFFEST option was used to determine if the unique invariance model had a significantly worse fit than Model 3 – the strong invariance model. Note, however, that while unique/strict invariance testing is generally considered to be too restrictive it was never-the-less conducted for demonstration purposes only. The model estimation terminated normally. While the chi-squared DIFF test value was significant ($\chi^2(9) = 64.01$, $p < .001$), due to a minimal change in the CFI (.001), the model passed the test of unique factorial invariance (refer to *Table A12*).

In conclusion, unique measurement invariance of the EF 1 measurement model that represents the working memory, planning, and problem-solving components of Executive Functioning was supported and this model became the new baseline model for further tests of the reliable latent-variable relationship.

Executive Functioning 2 (EF 2). The results across the four levels of measurement invariance tested for EF 2 (from configural, to weak, to strong, to unique) are nearly identical to the results of the four models described in detail above for EF 1:

there was a significant chi-square value ($p < .001$), yet a close-fitting CFI (<.98) with minimal change (<.002) based on the chi-square DIFF test (see *Table A12*). The WRMR was also a good fit to the data (.069 – .072). The only notable difference in the results across the four models for EF 2 compared to EF 1 is that for EF 2 the RMSEA had mediocre fit (rather than fair fit) across the models (.075 – .91).

In conclusion, unique measurement invariance of the EF 2 measurement model as represented by four indicators that represents the shifting and emotional regulation components of Executive Functioning was supported and this model became the new baseline model for further tests of the reliable latent-variable relationship.

Prosocial Skills. Similarly, the results across the four levels of measurement invariance tested for Prosocial Skills (from configural, to weak, to strong, to unique) are nearly identical to the results of the four models described in detail above for EF 1 and EF 2: there was a significant chi-square value ($p < .001$) for Models 3 and 4, yet a close-fitting CFI (>.98) with minimal change (<.003) based on the chi-square DIFF test (see *Table A12*). The RMSEA had fair model fit (ranging from .057 – .064) and the WRMR suggested the model was a good fit to the data across the four models of invariance tested. The only notable difference was that the chi-square DIFF test was not significant for Model 2 (weak invariance testing), but this difference is hardly worth mentioning as the change in CFI was used to determine whether to proceed with subsequent model invariance testing (in this case, strong invariance – Model 3).

In conclusion, unique measurement invariance of the Prosocial Skills measurement model as represented by five indicators was supported and this model became the new baseline model for further tests of the reliable latent-variable relationship.

Structural Model

Model Fit

The relationship between EF 1, EF 2, and Prosocial Skills on academic achievement are portrayed in Figure B2 (Math – Model 1) and Figure B3 (Reading – Model 2). The fit statistics for both models are presented in Table A13 and indicated the data fit the models well.

Within-Time Correlations

As presented in Table A14 the correlations among the non-academic variables of EF 1, EF 2, and Prosocial Skills were positive and significant, and the effect sizes, as defined by Cohen et al. (2003), were large (r = .71 – .92). The correlations among each pair of variables remained relatively stables across time from the fall of Kindergarten to the spring of Grade 2. The strength of the correlations among EF 1 and EF 2 becomes stronger in the spring of Kindergarten (increases from .78 to .89) and remains strong and relatively stable through to Grade 2 but is indicative of multicollinearity (r = .87 – .92). The correlation between EF 1 and Prosocial Skills and EF 2 and Prosocial Skills suggested minor fluctuations (r < .08) over time.

With regards to the correlation between these variables (EF 1, EF 2, and Prosocial Skills) and the academic achievement variables (Math and Reading), correlations were positive and mostly significant. Specifically, for the correlations of the non-academic variables with Math, the effect sizes were small for the 16 of the 18 correlations that were significant (r = .07 – .30), as were the effect sizes for all 18 correlations with Reading (r

= .06 – .22). Broadly, EF 1 had larger correlations with both academic variables than did EF 2, and Prosocial Skills had the smallest correlations with both academic variables.

Auto-Regressive Paths

Table A15 presents the completely standardized parameter estimates for the auto-regressive paths for both models of Math (Model 1) and Reading (Model 2). These paths, which are all statistically significant and meaningful, are shown in Figures B2 and B3 with heavy set solid lines. Meaningful is defined as consistent (in direction of sign) across multiple time points and had a standardized coefficient that was .16 or greater. Note that the results for the non-academic variables (EF 1, EF 2, and Prosocial Skills) are identical in both models.

The stability coefficients across EF 1 (ß = .64 – .83) and EF 2 (ß = .52 – 78) were positive, significant, and large, and as the range indicates, these values remained relatively stable over time. These results suggested good levels of measure stability over time and that a large portion of the variance in the variable at a later time point is explained by the same construct at the preceding time point.

This was not the case, however, for Prosocial Skills. While the stability coefficients were positive and significant over time, the within-grade (i.e., Kindergarten fall to spring) coefficients were quite large (ß = .90 – .92), but the coefficients that included the summer lag time (e.g., spring Kindergarten to fall Grade 1 and spring Grade 1 to fall Grade 2) were medium (ß = .31, .34, respectively).

The stability coefficients across the Math and Reading variables respectively were positive, significant, and medium for Math (ß = .43 – .57) and large for Reading (ß = .68 – .85). The value of the coefficients remained relatively stable over time for both models.

As with the other non-academic variables described above, the results suggested good levels of measure stability over time and that a small to large portion of the variance in the variable at a later time point is explained by the same variable at an earlier time point. However, contrary to the other variables in the model, the stability coefficients *increased* slightly in value between the spring and subsequent fall of each year.

Cross-Lagged Paths

Table A16 presents the completely standardized parameter estimates for the cross-lagged paths for both models of Math (Model 1) and Reading (Model 2). Note the model parameter estimates for the cross-lagged pathways between EF (EF 1 and EF 2) and Prosocial Skills are identical in both models, with the exception, of course, of the pathways between these variables and Math and Reading.

All significant cross-lagged pathways, however marginal the effect size, are discussed below and represented as solid lines in the models (refer to *Figures B2 and B3*). However, the statistically meaningful cross-lagged pathways are represented as heavy set solid lines in the models. Pathways deemed meaningful were consistent (in direction of sign) across multiple time points and had a standardized coefficient that was .16 or greater. Based on these criteria, the statistically meaningful cross-lagged pathways were as follows. With regards to Aim 1, there was a significant, positive, and bidirectional relationship between EF 2 and Prosocial Skills, but the magnitude of the pathways from EF 2 to Prosocial were deemed meaningful. With regards to Aim 2, EF 1 was a significant, positive, and meaningful predictor of both Math (stable across time) and Reading (declined across time).

Aim 1 Bidirectional Relationship. The statistically significant cross-lagged pathways among EF 1, EF 2, and Prosocial Skills are described below, which concern hypothesis 1 (H1). Additionally, results concerning hypothesis 3 (H3), children's gender effects, are outlined. Because no statistically significant results were obtained for hypothesis 2, these findings are not presented below.

EF 1 as a Predictor of Prosocial Skills (H1). With the exception of the spring of Kindergarten through the fall of Grade 1, EF 1, which represents the working memory, planning, and problem-solving components of EF, significantly and negatively predicted Prosocial Skills across the time points from the fall of Kindergarten through the spring of Grade 2. For a one SD increase in EF 1 (working memory, planning, problem solving) there was a marginal *decrease* in the SD of Prosocial Skills with this relationship remaining relatively constant (ß = -.11 – -.15).

Prosocial Skills as a Predictor of EF 1 (H1). From the spring of Kindergarten through the Spring of Grade 1 Prosocial Skills significantly, but inconsistently and marginally predictive of a student's levels of EF 1 (ß = -.06, .03).

EF 2 as a Predictor of Prosocial Skills (H1). EF 2, which represents the shifting and emotional regulation components of EF, significantly and positively predicted Prosocial Skills across the time points from the spring of Kindergarten through the spring of Grade 2. The coefficients ranged between .16 and .30, but this range excludes the coefficient that was significant, but not meaningful in Grade 1 (fall to spring, ß = .04, p = .02). Thus, for a one SD increase in EF 2, there was an increase in Prosocial Skills (ß = .24, .04, .30, .16). In other words, these findings suggested that having greater shifting and emotional regulation of symptoms yielded a higher score on Prosocial Skills, but the pattern of this relationship is somewhat unstable.

Prosocial Skills as a Predictor of EF 2 (H1). Prosocial Skills significantly, positively, and marginally predicted EF 2 within each grade (K- Gr.2), such that for a one SD increase in Prosocial Skills there was a small SD increase (ß = .07 – .12) in EF 2, as represented by shifting and emotional regulation. However, this relationship was not significant across the two time points that included the summer lag time (spring to fall of Kindergarten to Grade 1 and Grade 1 to Grade 2).

Children's Gender Effects (H3). When children's gender was regressed on EF 1, EF 2, and Prosocial Skills the standardized parameters (STDY) results suggested a small but statistically significant effect at one or more time points. For simplicity, Figures 2 and 3 do not reflect children's gender effects. Recall that children's gender was assessed using teacher reports (0 = *male*, 1 = *female*). Being identified as a girl was associated with a marginally lower score on EF 1, which measures working memory, planning, and problem solving, as compared to being identified as a boy in Grade 1 through Grade 2 (ß = - .06 – -.10, *SE* = .02 –- .04, *p* = .001 – .03). To give an example of this effect size, because the mean score of EF 1 for Cohort 1 and 2 ranged from 2.66 to 2.96 (on a Likert scale with a potential from 0 to 4), the SD represents at most a .30 decline in girls' scores on EF 1. Similarly, being a girl was associated with a marginally lower score on EF 2, which measures shifting and emotional regulation, compared to being a boy, across all time points from the spring of kindergarten through the spring of Grade 2 (ß = - .07 – .13, *SE* = .02 –- .04, *p* = .001 – .02). Surprisingly, girls also had a small but significant decrease in their standardized scores of Prosocial Skills across from the fall of Grade 1 through the spring of Grade 2 (ß = -.07 – -.18, *SE* = .03 –- .05, *p* < .001) compared to boys, but again this represents at most a .30 decline in girls' scores on the Prosocial Skills scale (e.g., Cohort 1 fall Grade 1 mean of 1.55 x -.18, see *Table A6*).

Being identified as a girl meant a significant small to moderate greater value in standardized Math scores as compared to boys (but not Reading). This was observed across all time points from the spring of Kindergarten through the spring of Grade 2 (ß = .11 – 26, *SE* = .04 – .06, *p* < .001 – .003). There was a slight u-shaped curve observed such that the value of the standardized scores decreased starting in the fall of Grade 1 (from .26 to .13) and then increased at the final time point observed in the spring of Grade 2 (from .11 to .17).

Aim 2: Academic Achievement. The statistically significant cross-lagged paths between each of EF 1, EF 2, and Prosocial Skills with academic achievement (Reading and Math) are described below, which concern hypothesis 4 (H4) and hypothesis 5 (H5).

EF 1 as a Predictor of Math (H4). EF 1 significantly and positively predicted Math across the time points from the fall of Kindergarten through the spring of Grade 2, with a medium size effect observed (ß = .24 – .42), suggesting that for one SD increase in EF 1 (working memory, plan, problem solving) there was a .24 to .42 SD increase in Math scores. An effect size in this range is meaningful as recall Math score represent percentage correct. For example, for Cohort 1 this represents an 8.8% increase in Kindergarten spring Math scores (21.01 x .42, see *Table A7*, Cohort 1, T2). Examining effect sizes across time, the magnitude of this association slowly decreased, then increased such that from spring of Kindergarten to the spring of Grade 1 the coefficient declined slightly (from .42 to .29 to .24), then rose slightly (to .32 and .31) from the fall to the spring of Grade 2 but remained positive in direction and moderate in strength.

EF 1 as a Predictor of Reading (H4). EF 1 also significantly and positively predicted Reading from the fall of Kindergarten through the fall of Grade 2. The effect size initially observed was medium (ß = .38, *SE* = .4, $p < .001$), which means it represented an increase in about 13 words per minute for Cohort 1 (33.70 x .38, see *Table A7*, Cohort 1, T2).

However, the effect size of this relationship halved from the spring of Kindergarten through the fall of Grade 1. Thus, as students moved forward in grade level the relationship between EF 1 and Reading became not particularly meaningful by the fall of Grade 2 (ß =.13, *SE* = .03, $p < .001$) and became non-significant for the final time point when assessed in the spring of Grade 2.

EF 2 as a Predictor of Math (H4). EF 2 (shifting, emotional regulation) at the preceding time point significantly and negatively predicted student math scores across three of the five time points observed. In the fall of Kindergarten for a one SD increase in EF (i.e., better emotional regulation and shifting) there was a .11 SD *decrease* in Math achievement (*SE* = .04, $p < .01$) the following spring (Kindergarten). This represented a 2.3% decrease in math scores. A similar relationship was observed at two later time points from the spring of Grade 1 through the spring of Grade 2 (ß = -.13, -.12). Given the small effect size in conjunction with an inconsistent relationship across time these results are not very meaningful.

EF 2 as a Predictor of Reading (H4). EF 2 in the fall of Kindergarten was significantly and positively predicted students reading scores in the spring of

Kindergarten (ß =.16, *SE* = .02, *p* <.001), such that a one SD in EF 2 represented students being able to read an additional five words per minute (33.70 x .16, see *Table A7*, Cohort 1, T2). This relationship was also significant, although even less meaningful, in the fall of Grade 1 on students reading scores in the spring of Grade 1 (ß =.06, *SE* = .02, *p* <.001). Given that this relationship was inconsistent over time and not very strong to begin with (in Kindergarten), these results are also not very meaningful.

Prosocial Skills as a Predictor of Reading (H5). Prosocial Skills also significantly predicted Reading (but not Math) from the spring of Kindergarten through the fall of Grade 1. In the fall of Kindergarten for a one SD increase in Prosocial Skills there was a .07 SD *decrease* in Reading (*SE* = .03, *p* = .05) the following spring (Kindergarten). A similar coefficient was observed for the Spring of Kindergarten to the Fall of Grade 1 (ß = -.06, *SE* = .02, *p* = .04). While these effects were statistically significant, and in an unexpected direction (given the raw correlations), they are not meaningful given they represents, for example, a decrease in roughly two words read per minute for Cohort 1's spring Reading scores (33.70 x .07, see *Table A7, Cohort 1, T2*). Furthermore, the effect size of this relationship became non-significant over time.

DISCUSSION

Second Step® is the most widely distributed Social-Emotional Learning (SEL) classroom curriculum used in the United States. The program was revised in the last decade to explicitly target executive functioning (EF) skills in conjunction with social-emotional skills. EF skills are broadly defined as the cognitive functions that enhance goal-related problem-solving behavior. Second Step® was revised on the premise that the integration of these two domains (SEL Skills and EF) may bolster academic improvements. Despite this, studies to date with the revised Second Step® have yielded inconsistent and weak effects on reading and math achievement (e.g., Low et al., 2019), but only direct effects have been examined. In theory, revised program logic would hypothesize impacts on academic achievement (distal outcome) via improvements in SEL Skills and/or EF. Although the current study does not examine intervention effects, its aims delineate the reciprocal, longitudinal relationship between SEL Skills, EF, and academic impacts; an important validation of underlying program logic/theory for Second Step® and SEL program more broadly.

The current study had two general aims: the bidirectional longitudinal relationship between EF and Prosocial Skills (as a proxy for SEL Skills); and the unique and co-joint relationship of EF and Prosocial Skills to Reading and Math achievement.

Aim 1: Bidirectional Relationship

The first major aim of this study, which was to examine the bidirectional relationship between EF and Prosocial Skills, involved two hypotheses, which will be discussed sequentially.

EF 2 Predicts Prosocial Skills (H1)

First, to examine H1, which was that *EF and Prosocial Skills will positively and reciprocally influence each other across time points, controlling for the stability of each construct respectively.* The current study's data lends some supports this hypothesis, but this needs to be qualified.

There was support for a bidirectional relationship between EF 2 (shifting, emotional regulation) and Prosocial Skills. In the case of Prosocial Skills contributing to the development of EF 2 (shifting/emotional regulation), those relationships, while significant, were marginal, less consistent, and limited to within-school year associations. Such a trend is interesting and may suggest that classroom specific social interactions bolster related skills of shifting and emotional regulation. The school day is filled with transitions to various activities within and outside of the classroom (e.g., different workstations, break times, physical education, library time) and while the timing of some of these activities is very consistent (i.e., break times), the timing of others is more fluid. The data suggests children with more Prosocial Skills have a slightly easier time with these transitions while these specific activities are in effect during the school year. Perhaps being able to shift one's perspective more routinely focused outside of oneself onto the other is a strength in the classroom because it broadens one's "visual field", so to speak, during times of transition that act as a cue or reminder that transition times are not self-specific nor self-isolated acts but are more like a community event. It is important to remember, however, that this relationship is marginally meaningful.

However, the strength of the relationships between EF 2 and Prosocial Skills across time were consistent and much stronger, and thus indicative of a greater cascade effect, than the relationship between Prosocial Skills and EF 2. The data that EF 2 predicts Prosocial Skills lends support to the assumption that some innate hard-wired skills (i.e., shifting, emotional regulation) may temporally precede the acquisition of socially driven, interactive Prosocial Skills that may require not only underlying foundational cognitive capabilities, but time to develop as children learn to function in various social environments. From Piaget's theory of human development, I could posit that classroom environments, like all environments, are predicated on the ability of the organisms (children in this case) to adapt to change, be it a change in classroom daily routines, transitioning from one subject to the next, and changes in the emotional temperature of a classroom throughout the day. However, unlike the case of an ameba, in the case of the human child, that environment is a social environment. The ability to respond to this change may function to free up energy that would otherwise be spent regulating oneself (i.e., emotional regulation) in a changing environment (i.e., shifting) and in turn being able to look outside of oneself to one's peers. Classrooms, much like human society depends on cooperative, prosocial behavior. A child who is transitioning better will have more time to engage with peers during these times of transition. Furthermore, recall that it is teachers who are rating both of these skills (EF and Prosocial Skills). It is possible that children who adjust well to changes in plans and can better regulate their emotions may be perceived by teachers as more able to meet the demands of their social environment, and thus, be deemed more prosocial.

That being said, the relationships between EF 2 and Prosocial Skills should be qualified herein with regards to two pathways. First, this relationship is not significant in Kindergarten (fall to spring pathway). Kindergarten is a time when children's executive functions are still more unitary in nature and thought to be less compartmentalized due to the stage of their brain maturation (e.g., Best et al., 2009; Peterson & Welsh, 2014; Wiebe et al., 2011). Brain maturation then may take a little bit of time to catch up such that it is an adaptive function enabling kids to adapt to the social demands of the classroom by the spring. Second, the relationship emerges as meaningful and significant from the spring of Kindergarten to the fall of Grade 1, but then remains significant (but not meaningful) in Grade 1 (fall to spring). In examining one slice of this pie (i.e., shifting), perhaps Grade 1 is a transitional time in the development of children's shifting abilities. While we know that shifting develops linearly and progressively (Best & Miller, 2010), Anderson's (2002) review of nearly a dozen studies suggested shifting undergoes a rapid period during middle childhood (ages 7-9). Thus, for the component of shifting, at least, perhaps there is an emerging – receding – emerging element in development of children's shifting that may in turn be contributing to the initially inconsistent relationship between EF 2 and Prosocial Skills in the model.

Progressive Reciprocal Association Not Supported (H2)

Hypothesis 2 predicted that *the magnitude of the reciprocal association between EF and Prosocial Skills (i.e., the cross-lagged pathways) would strengthen over time*. This hypothesis was predicated on an assumption based on two notable trends in the research: first, that EF and Prosocial Skills increase in normally developing children, and second, that these two constructs are generally found to be at least moderately correlated. While I did not find a progressive relationship in the sample of early elementary school aged children in the current study, data with a population that spans a larger age range may yield different results. Middle school is a time when children transition away from solitary work assignments to more social cooperative assignments. In these environments, peers with slightly greater cognitive abilities would be in a position to scaffold their peers, possibly contributing to gains in cognitive growth of executive functions. In other words, because EF develops at slightly different rates of cognitive expansion (e.g., myelination, neuron proliferation, synaptogenesis) and regression (e.g., cell death, synaptic pruning) even among same-sage peers (Best & Miller, 2010), peer and friendship dyads could be viewed as dynamics in which scaffolding can occur. The child is also an active participant in the social skills they are creating with their peers (i.e., "social agent"; Vygotsky, 1930-1934/1978) and it has long been established social skills such as prosocial behavior are fostered in the peer and friendship dynamic (e.g., Cillessen et al., 2005; Meuwese et al., 2016; Poorthuis et al., 2012). It is those same social skills that are developing in the context of this relationship dynamic that has the potential to foster cognitive development (Hartup, 1996). In other words, Prosocial Skills that foster learning in the zone of proximal development can create the environment for executive function growth in middle elementary school when the environment better provides the opportunity for such an exchange.

Marginal Gender Effects (H3)

In Hypothesis 3 I further proposed that *children's gender will influence EF and Prosocial Skills*. While there is support for this hypothesis, the effects are marginal. In the current study girls had significantly lower scores as compared to boys on all components of EF (working memory, planning, problem solving, shifting, emotional regulation). Broadly, research suggests that the effects of gender on the development of EF are inconsistent (e.g., see Brocki & Bohlin, 2004) and can vary based on type of measure used to collect data (Fournet et al., 2015; Peterson & Welsh, 2014), with rating scale results not well studied (Catale et al., 2015). Results can also vary by type of informant (Thorell et al., 2013).

Furthermore, while the current study found girls had significantly lower Prosocial Skills compared to boys, given the ample research to suggest otherwise (see Eisenberg & Fabes, 1998; also see Adams et al., 1999; Caputi et al., 2012; Carlo, 2014; Côté et al., 2002; Knafo & Plomin, 2006; Scourfield et al., 2004; Zimmer-Gembeck et al., 2005), and that scores were only marginally lower herein, it would be premature to make inferences based on these results.

Aim 2: Executive Functioning and Prosocial Skills Predict Academic Achievement

The second aim of this study, which concerns hypotheses 3 and 4 below, was to examine the longitudinal relationship of both EF and Prosocial Skills to children's reading and math achievement.

EF 1 is a Stronger Predictor than EF 2 (H4)

The results of the current study provide support for Hypothesis 4 that predicted EF (1 and 2) would directly influence academic achievement. While there is support for this hypothesis, this statement must be qualified in two ways.

First, EF 1 is a stronger predictor of Math than Reading. The longitudinal study of EF and academic achievement in school aged children using EF tests typically measures the combination of working memory, inhibition, and shifting due to the popularity of the associated well-established EF tests (refer to *Table A1* and *Appendix C*), rather than the single component of working memory or planning/problem solving (Baggetta & Alexander, 2016; Best et al., 2009). As such, there is a comparatively limited amount of research to which to compare the results of the current study, in which EF 1 was a consistent and moderate predictor of Math from Kindergarten through Grade 2. Similarly, research examining Math as an outcome variable suggests small to medium size effects with both EF test mesuring working memory (e.g., see Adams & Hitch, 1997; Cragg & Gilmore, 2014; De Smedt et al., 2009; Friso-Van den Bos et al., 2013; Geary, 2011) and planning/problem solving (e.g., Friedman et al., 2014), as well as the combination of these components assessed by EF rating scales (e.g., Decker, 2017; Thorell & Nyberg, 2008; Thorell et al., 2013; Waber et al., 2006). Reading, on the other hand, is less commonly examined in relation to EF in general, and working memory, planning/problem solving specifically (e.g., Thorell et al., 2013), and in this study, as has been found elsewhere (e.g., Geary, 2011), working memory and planning/problem solving components of EF appear to be more salient predictors of reading for younger students (e.g., Kindergarten and Grade 1). When children are first learning to read, they need to remember a combination of sounds and to read in a particular order (e.g., right to left, up/down, space between words). However, as children advance in grade level, these particular reading skills, which are closely associated with working memory (refer to *Table A2*), do not require additional development. Math, on the other hand, is essentially a problem solving skill that requires the use of working memory (e.g., Bull & Lee, 2014; Goeff et al., 2005) with rules becoming progressively more complex, unlike the rules for reading that do not change. Children in early elementary school are learning computations (addition, subtraction, basic multiplication and division) and order of computations (i.e., BEDMAS), which requires them to preform steps of a task in order (working memory) and to use resources (i.e., bead counting) to solve a problem (planning/ problem solving). Thus, the results of this study lend support for the idea that the proportion of variance in a child's reading skill is largely accounted for by one's previous reading ability, particularly as children transition out of Kindergarten and onwards, which is not the case for math (refer to *Table A15*).

The second caveat of hypothesis 4 is that EF 1 is a stronger predictor than EF 2 (shifting/ emotional regulation) of academic achievement (and more precisely, of math, as described above). As previously discussed, our understanding of emotional regulation as a predictor of children's academic achievement has seldom been studied using EF tests, with EF rating scales generally lacking in longitudinal investigation (refer to section *Executive Functioning as a Predictor of Academic Achievement*). Meta-analyses of the relationship between EF tests of shifting and reading and math suggest a weak correlation (Yienad et al., 2013), however, shifting is a common component of EF tests based in a multi-factorial component of EF of working memory, inhibition, and shifting, with findings related to academic achievement likely attributed to working memory (Bull & Lee, 2014). Such results are in line with the current study in which EF 2 (shifting/ emotional regulation) was not a significantly meaningful longitudinal predictor of Math or Reading. Shifting and emotional regulation likely have greater implications for non-academic related performance, and as discussed earlier (refer to section *EF 2 Predicts Prosocial Skills (H1)*), one of these implications may be the development of Prosocial Skills.

EF 1 is a Stronger Predictor than Prosocial Skills (H5)

Hypothesis 5 proposed that EF would be a stronger predictor of academic achievement than Prosocial Skills. This hypothesis is largely supported but must be qualified in light of the finding that EF 1 (working memory, plan/problem solving) is a stronger and hence more meaningful predictor of academic achievement than EF 2 (shifting/emotional regulation).

In the current study, Prosocial Skills was not a significant predictor of Math, nor a meaningful predictor of Reading for early elementary school children (refer to section *Results – Prosocial Skills as a Predictor of Reading*). Contrarily meta-analytic studies suggest there is a large body of evidence that SEL Skills are related to school outcomes (e.g., Durlak et al., 2011; Taylor et al., 2017). An important caveat made by Durlak et al. (2011) is that this relationship is stronger, and therefore more meaningful, when a combination of SEL Skills is measured. The current study only examined one SEL Skill, that of Prosocial Skills, because the most commonly cited measures of SEL Skills (i.e., SSIS; Gresham & Elliott, 2008; SSIS SEL – RF; Gresham & Elliott, 2017) are in fact a measure of prosocial behavior (refer to section Major Dimensions of SEL Skills and *Appendix C*), aligning with the CASEL Core-5 (Elliott et al., 2017; Gresham et al., 2018). That being said, the measure of Prosocial Skills used in the current study was not as comprehensive as, for example, rating scales such as the SSIS SEL – RF (Gresham & Elliott, 2017). As such, the results of the current study reiterate Durlak et al.'s notion that a number of SEL Skill need to be measured in order to find an achievement effect.

However, one noteworthy benefit of the current study is that I directly compared the effects of Prosocial Skills versus EF on academic achievement in the same model, with results suggesting that EF (i.e., EF 1), which was merely a 9-item measure of working memory, planning, and problem solving, is a much stronger predictor of academic achievement. Taking all this in consideration with the number of SEL Skills that need to be measured to find an achievement effect, one conclusion offered herein is that EF has been seriously undervalued in the preponderance of SEL Skills intervention worldwide. Not only are student gains in reading and math for a smooth transition through the grades a major aim of American schools (e.g., Common Core Standards Initiative, 2012), but is a goal shared by schools worldwide. Because EF is a developmental skill that crosses cultural boundaries (e.g., Thorell et al., 2013) and SEL programs only sometimes include ways to bolster EF, interventions that bolster EF could have widespread implications for global administration of school-based interventions whose goal is to find an intervention effect (e.g., bolster academic achievement).

Furthermore, interventions that bolster EF in a normative population of children provide support for those same children who are at risk of developing EF deficits compared to their peers. Implications for children who have difficulty with daily living due to EF deficits but who do not meet the diagnosis for a clinical EF syndrome nor have specific cognitive deficit (e.g., language, attention, motor) have an impairment termed dysexecutive functioning. Cognitive deficits of dysexecutive functioning include reduced working memory, planning and organizational problems, difficulties generating and/or implementing strategies, and poor utilization of feedback (Gioia et al., 2000; 2001), which comprises EF 1. Thus, including measuring children's EF in in early elementary school could additionally provide a screening mechanism in normative populations to minimize long-term potential for children not achieving their full academic potential in core curriculum subjects like reading and math, not to mention the psycho-social impact of having difficulty engaging in routine tasks that are seemingly second nature to those with greater EF capacities.

As discussed, some measures of SEL Skills have begun to incorporate EF items. Recall the DESSA- SSE (Devereux Center for Resilient Children, 2012) has a 36% overlap in its items with established childhood rating scales of EF (refer to section *Conceptual Relationship Between Executive Functioning and SEL Skills*), with 16 of 36 of these items parallel commonly cited rating scales of EF's components of working memory, planning, and problem solving. However, as previously argued, for conceptual clarity it is important that such measures acknowledge they are measuring EF skills, rather than labeling such skills as SEL Skills.

Limitations

The current study's limitations are twofold. First, due to secondary data use, a measure of EF was created herein. Its two-factor solution containing four of the five most commonly researched components of EF (see Baggetta & Alexander, 2016), and is psychometrically on par with other reputable scales (i.e., CHEXI, Thorell & Nyberg, 2008). However, highly developed scales have the advantage of better distinction among the factors, unlike the current study's measure of EF that contained two correlated factors that were deemed moderately (r = .66, based on EFA results) to highly (r = .78 – .92, based on structural model results) correlated, bordering on multicollinearity. That being said, the intercorrelations among the factors of the most commonly cited rating scales have not been published in factor analyses studies of Thorell and Nyberg's (2008) widely accessible CHEXI (Camerota et al., 2018; Catale et al, 2012, 2015, 2018; Thorell & Nyberg, 2008; Thorell et al., 2010). As for the BRIEF (Gioia et al., 2000), while Lyons et al. (2016) report the intercorrelation among the two factors (the *MCI* and *BRI* indices) is moderate (r = .50, based on EFA results), there is some controversy on both the specified two factor solution (e.g., Egeland & Fallmyr, 2010; Fournet et al., 2015; Jimenez & Lucas-Molina, 2019) and its latest version's (BRIEF-2; Gioia et al., 2015) three factor solution (for a review see Jimenez & Lucas-Molina, 2019).

To date, the most widely utilized rating scales of EF have had very limited factorial invariance testing (for a review see Jimenez & Lucas-Molina, 2019), although there has been some success reported with a limited version (problem solving, attention, behavioral/emotional regulation) of Gioia et al.'s (2000) BRIEF (see Garcia-Barrera et al., 2011; 2013). The current study's measure of EF, on the other hand, demonstrated sound longitudinal invariance.

Measurement limitations aside, the current study's findings cannot be generalized beyond early elementary school aged children. This is an important consideration, as curriculum, classroom demands, and teacher-student relationship transform considerably across one's school trajectory. Thus, the extent to which each skill (i.e., EF, Prosocial) contributes to the development of the other may change or strengthen as children's schoolwork becomes more integrated and partner oriented. In brief, the current findings should be viewed in light of the developmental period (early elementary) during which the study took place.

Implications

Schools around the globe purchase SEL interventions posited as being salient direct or indirect predictors of children's academic achievement, yet meta-analyses indicate effect sizes are small (Durlak et al., 2011), and often based on non-core academic indicators (e.g., percent meeting minimum learning benchmarks). In the case of the meta-analysis by Taylor, Oberle, Durlak & Weissburg (2017), the authors noted academic outcome effects sizes averaged .33; however, only 8 published studies were found that examined academic outcomes, and most of these programs were multi-tiered, intensive family-school or wrap-around models, and many of the largest studies to date were not included. Some of these interventions function to hone both EF and SEL skills (Bierman & Torres, 2016), such as recent changes to Second Step®, with a modified rating scale (i.e., DESSA – SSE; Devereaux Center for Resilient Children, 2012) that includes EF-type items. However, DESSA – SSE is primarily a SEL skills scale used to measure intervention effects between SEL Skills and academic achievement. As previously discussed, SEL Skills items from the DESSA – SSE highly overlap with items that are thought to measure EF components of working memory, inhibition, shifting, planning/problem solving, and emotional regulation (refer to section *Construct Overlap: Ratings Scales Item Similarities*). Therefore, it is important for researchers to realize that when they are using this scale that they are measuring EF components, and in turn not to misattribute results of a relationship between SEL skills and academic achievement, when in fact, EF is being measured.

Furthermore, it is important to be able to determine which components of EF and SEL Skills affect which outcomes. One reason SEL interventions may not be as effective as intended is that they do not always include ways to bolster EF, and when they do, the program may not be designed to ensure the EF and SEL skills are being taught in a way that supports their mutual development in relation to each other and to academic achievement. For example, there is a tendency in EF and SEL Skills interventions to predominantly target EF components of working memory, inhibition, and shifting (e.g., Head Start Research-Based Developmentally Informed (REDI), Bierman et al., 2008a;

Röthlisberger et al., 2012), with marginally significant improvement in said EF components (e.g., see Bierman et al., 2008b). Yet as the results of the current study suggest, this combination of components of EF may not be optimal when attempting to bolster both SEL Skill outcomes and academic achievement.

The purpose of this paper was an attempt to hone the direction of SEL interventions in two ways. First, by exploring which, if any, of the EF components are most likely to positively influence the development of SEL Skills. This exploration began by resuming an overlooked discussion in the fields of EF and SEL Skills on the conceptual and empirical relationship between these two constructs. Given the results of the current study suggest that these hard-wired EF skills of shifting and emotional regulation influence the development of Prosocial Skills, the implication is that by incorporating ways to bolster children's EF's components of shifting and emotional regulation in SEL programs may in turn bolster the very social skills, namely prosocial behavior, that these programs were intended to foster. It is posited that this may in turn strengthen program outcomes.

Second, given that SEL interventions aim to improve academic achievement, the current study informs us that EF components of working memory and planning/problem solving are the most salient. Currently interventions that bolster the EF component of planning/problem solving are scarce, while interventions that include working memory typically involve video games, sociodramatic play, or small groups (Bierman & Torres, 2016; Diamond & Lee, 2011), with some promising results (see Kroesberg et al., 2014; St Clair-Thompson et al., 2010; Thorell et al., 2009; Traverso et al., 2015). However, these promising results do not always extend to improved academic achievement (Bierman &

Torres, 2016; Cragg & Chevalier, 2011; St Clair-Thompson et al., 2010), and as such EF interventions have challenges to overcome regarding their efficacy (Bierman & Torres, 2016; Morrison & Chein, 2011). In addition to bolstering EF components, this paper helps make the case for isolating, labeling and discreetly measuring the program inputs in order to optimize evaluations of program validity.

CONCLUSION

To date, Second Step®, which is the most widely distributed SEL intervention program in the US, has found no program effects related to academic achievement. This raises questions about the relationship between SEL Skills and academic achievement more broadly, both within and outside of the context of interventions. While meta-analyses have been conducted suggesting a relationship between SEL Skills and academic achievement (e.g., Durlak et al., 2011; Taylor et al., 2017), this evidence does not appear as strong as evidence of the relationship between EF and academic achievement. Rarely have both constructs been tested longitudinally in one model in elementary school aged children. Additionally, some of the SEL literature, namely research and rating scales (e.g., DESSA – SSE), blurs the boundaries between these two constructs. While there is some overlap between these closely related (i.e., highly correlated) broad constructs, they come from distinct theoretical traditions with different definitions and measures. Ensuring that these measurement of these constructs remains distinct will in turn clarify the contribution of each construct to academic achievement.

In the current study, I first conducted a thorough review of the Second Step® measure, the DESSA – SSE (Devereux Center for Resilient Children, 2012) to examine its content overlap with widely cited measures of EF. Having identified the items with the greatest overlap (e.g., grounded in EF theory and measurement), I established longitudinal measurement invariance for a two-factor EF measure of working memory, planning, and problem solving (EF 1), and shifting and emotional regulation (EF 2). I was then able to test two overarching and related questions concerning EF and SEL. First, what is the longitudinal bidirectional relationship between EF (1 and 2) and SEL, using

Prosocial Skills as a proxy (Aim 1)? Second, of these two constructs (i.e., EF and Prosocial Skills), which is the most salient in predicting academic achievement (Aim 2)?

The results of the current study suggested that there is a bidirectional relationship between EF 2 (shifting/ emotional regulation) and Prosocial Skills, but that the strength of this association is strongest in the direction of EF 2 predicting Prosocial Skills. The data lends support to the assumption that innate hard-wired skills (i.e., EF) would temporally precede the acquisition of socially driven, interactive Prosocial Skills that may require not only underlying foundational cognitive capabilities, but time to develop as children learn to function in various social environments. Children who transition better throughout the day in the micro-evolving nature of the social environment that is their classroom may be in turn better able to engage with peers during these times of transition, and hence, behave more prosocially.

The results of the current study also suggest that the EF components of working memory, planning, and problem solving (EF 1) is a more significantly meaningful predictor of academic achievement, particularly Math, compared to EF 2 (shifting and emotional regulation) and Prosocial Skills. The finding that EF 1 is a stronger predictor of academic achievement than EF 2 is consistent with the literature and suggests shifting and emotional regulation likely have greater implications for non-academic related performance, such as the development of Prosocial Skills. Furthermore, because EF 1 is a stronger predictor of Prosocial Skills and given that only some SEL skill interventions incorporate EF, these results suggest that SEL programs that want to find an achievement effect should consider incorporating ways to bolster EF components of working memory, planning, and problem solving. Beyond the scope of academic achievement, interventions that enable children's social skills and executive functions to thrive will ultimately provide the fertile ground for the next generation of cooperative problem solvers to emerge.

REFERENCES

Adams, J. W., & Hitch, G. J. (1997). Working memory and children's mental addition. *Journal of Experimental Child Psychology, 67*(1), 21-38. doi:10.1006/jecp.1997.2397

Adams, J. W., Snowling, M. J., Hennessy, S. M., & Kind, P. (1999). Problems of behaviour, reading and arithmetic: Assessments of comorbidity using the strengths and difficulties questionnaire. *The British Journal of Educational Psychology, 69*(4), 571-585. doi:10.1348/000709999157905

Ahonniska, J., Ahonen, T., Aro, T., Tolvanen, A., & Lyytinen, H. (2000). Repeated assessment of the Tower of Hanoi Test: Reliability and age effects. *Assessment, 7*(3), 297-310. doi:10.1177/107319110000700308

Akhutina, T. V., & Pylaeva, N. M. (2011). L. Vygotsky, A. Luria and developmental neuropsychology. *Psychology in Russia: State of the art, 4*(1), *155-175.*

Alderman, N., Burgess, P.W., Knight, C., & Henman, C. (2003). Ecological validity of a simplified version of the Multiple Errands Shopping Tests. *Journal of the International Neuropsychological Society, 9*(1), 31-44. doi:10.1017/S1355617703910046

Allen, D. N., Strauss, G. P., Donohue, B., & van Kammen, D. P. (2007). Factor analytic support for social cognition as a separable cognitive domain in schizophrenia. *Schizophrenia Research, 93*(1), 325- 333. doi:10.1016/j.schres.2007.02.008

Alloway, T. P. (2007). Automated working memory assessment. London: Pearson Assessment. Translated and reproduced by permission of Pearson Assessment.

Alloway, T. P., Gathercole, S. E., & Kirkwood, H. J. (2008). *Working Memory Rating Scale (WMRS)*. London, England: Pearson Assessment.

Alloway, T. P., Gathercole, S. E., & Peckering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separate? *Child Development, 77*(6), 1698-1715. doi:10.1111/j.1467-8624.2006.00968.x

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology, 8*(2), 71–82. doi:10.1076/chin.8.2.71.8724

Anderson, V., Beauchamp, M. H., Yeates, K. O., Crossley, L., Hearps, S. J. C., & Catroppa, C. (2013). Social competence at 6 months following childhood traumatic brain injury. *Journal of the International Neuropsychological Society, 19*(5), 539-550. doi:10.1017/S1355617712001543

Andersson, U. (2008). Working memory as a predictor of written arithmetical skills in children: The importance of central executive functions. *The British Journal of Educational Psychology, 78*(2), 181-203. doi:10.1348/000709907X209854

Ansari, A., & Gershoff, E. (2015). Learning-related social skills as a mediator between teacher instruction and child achievement in head start: Learning-related social skills and teacher instruction. *Social Development, 24*(4), 699-715. doi:10.1111/sode.12124

Achenbach, T., McConaughy, S., Howell, C., & Masters, John C. (1987). Child/Adolescent Behavioral and Emotional Problems: Implications of Cross-Informant Correlations for Situational Specificity. *Psychological Bulletin, 101*(2), 213-232.

Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology, 5*(2), 115-129. doi:10.1076/chin.5.2.115.3167

Baddeley, A. D. (1986). *Working memory*. New York, NY. Clarendon Press.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*(1), 1-29. doi:10.1146/annurev-psych-120710-100422

Baddeley, A., Della Sala, S., Papagno, C., & Spinnler, H. (1997). Dual-task performance in dysexecutive and nondysexecutive patients with a frontal lesion. *Neuropsychology, 11*(2), 187-194. doi:10.1037/0894-4105.11.2.187

Baddeley, A., Emslie, H., Kolodny, J., & Duncan, J. (1998). Random generation and the executive control of working memory. *The Quarterly Journal of Experimental Psychology. A Human Experimental Psychology, 51*(4), 819-852. doi:10.1080/027249898391413

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47 – 90). New York, NY: Academic Press.

Baggetta, P., & Alexander, P. A. (2016). Conceptualization and operationalization of executive function. *Mind, Brain, and Education, 10*(1), 10-33. doi:10.1111/mbe.12100

Banfield, J. F., Wyland, C. L., Macrae, C. N., Munte, T. F., & Heatherton, T. F. (2004). The cognitive neuroscience of self-regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 62 – 83). New York, NY: Guilford Press.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*(1), 65-94. doi:10.1037//0033-2909.121.1.65

Barkley, R. A. (2001). The executive functions and self-regulation: An evolutionary neuropsychological perspective. *Neuropsychology Review, 11*(1), 1-29. doi:10.1023/A:1009085417776

Barkley, R. A., & Fischer, M. (2011). Predicting impairment in major life activities and occupational functioning in hyperactive children as adults: Self-reported executive function (EF) deficits versus EF tests. *Developmental Neuropsychology, 36*(2), 137-161. doi:10.1080/87565641.2010.54987

Barkley, R. A., & Murphy, K. R. (2011). The nature of executive function (EF) deficits in daily life activities in adults with ADHD and their relationship to performance on EF tests. *Journal of Psychopathology and Behavioral Assessment, 33*(2),137-158. doi:10.1007/s10862-011-9217-x

Barkley, R. A., (2012a). *Executive functions: What they are, how they work, and why they evolved*. New York, NY: Guilford Press. Retrieved from http://www.myilibrary.com.ezproxy1.lib. asu.edu?ID=357913

Barkley, R.A. (2012b). *Barkley deficits in executive function scale: Children and adolescents (BDEFS – CA).* New York, NY: Guildford Press.

Bateson, C.D. (2009). These things called empathy: Eight related but distinct phenomena. In J. Decety and W. Ickes (Eds.), *The social neuroscience of empathy* (pp.3-15). Cambridge, MA: MIT Press.

Batson, C. D. (1998) Altruism and prosocial behavior. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 2, pp.282-316, 4th ed.). Boston: McGraw-Hill.

Battistich, V., Solomon, J., Solomon, D., Watson, M., & Schaps, E. (1989). Effects of an elementary school program to enhance prosocial behavior on children's cognitive-social problem-solving skills and strategies. *Journal of Applied Developmental Psychology, 10*(2), 147-169. doi:10.1016/0193-3973(89)90002-6

Baughman, F., & Cooper, R. (2007). Inhibition and young children's performance on the tower of London task. *Cognitive Systems Research, 8*(3), 216-226. doi:10.1016/j.cogsys.2007.06.004

Beauchamp, M. H., & Anderson, V. (2010). SOCIAL: An integrative framework for the development of social skills. *Psychological Bulletin, 136*(1), 39-64. doi:10.1037/a0017768

Benenson, J. F., Markovits, H., Roy, R., & Denko, P. (2003). Behavioural rules underlying learning to share: Effects of development and context. *International Journal of Behavioral Development, 27*(2), 116-121. doi:10.1080/01650250244000119

Berg, E. (1948). A simple objective technique for measuring flexibility in thinking. *The Journal of General Psychology, 39*, 15-22. Retrieved from: https://search.lib.asu.edu/primo-explore/fulldisplay?docid =TN_medline18889466&context=PC&vid=01ASU&search_scope=Everything&tab=default_tab&lan g=en_US

Bergin, C. (2014). Educating students to be prosocial at school. In L.M. Padilla-Walker & G. Carlo (Eds.) *Prosocial development: A multidimensional approach* (pp. 279-302). Oxford Scholarship Online. doi:10.1093/acprof:oso/9780199964772.003.0014

Berlin, L., & Bohlin, G. (2002). Response inhibition, hyperactivity, and conduct problems among preschool children. *Journal of Clinical Child & Adolescent Psychology, 31*(2), 242-251. doi:10.1207/S15374424JCCP3102_09

Best, J., & Miller, P. (2010). A developmental perspective on executive function. *Child Development, 81*(6), 1641-1660. doi:10.1111/j.1467-8624.2010.01499.x

Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review, 29*(3), 180-200. doi:10.1016/j.dr.2009.05.002

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development, 70*(3), 636-644. doi:10.1111/1467-8624.00046

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science, 7*(3), 325-339. doi:10.1111/j.1467-7687.2004.00351.x

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge, UK: Cambridge University Press. Retrieved from http://www.ebrary.com.ezproxy1.lib.asu.edu

Bierman, K. L., Coie, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., McMahon, R. J., . . . Conduct Problems Prevention Research Group. (2010). The effects of a multiyear universal social-emotional learning program: The role of student and school characteristics. *Journal of Consulting and Clinical Psychology*, *78*(2), 156 - 168. doi:10.1037/a0018607

Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., . . . Gill, S. (2008a). Promoting academic and social-emotional school readiness: The head start REDI program. *Child Development, 79*(6), 1802-1817. doi:10.1111/j.1467-8624.2008.01227.x

Bierman, K. L., & Erath, S. A. (2006). Promoting social competence in early childhood: Classroom curricula and social skills coaching programs. In K. McCartney & D. Phillips (Eds.), *Blackwell Handbook on Early Childhood Development* (pp. 595-615). Malden, MA: Blackwell. doi: 10.1002/9780470757703.ch29

Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., & Domitrovich, C. E. (2008b). Executive functions and school readiness intervention: Impact, moderation, and mediation in the head start REDI program. *Development and Psychopathology, 20*(3), 821-843. doi:10.1017/S0954579408000394

Bierman, K.L., & Torres, M. (2016). Promoting the development of executive functions through early education and prevention programs. In J. A. Griffin, L. S. Freund, & P. McCardle (Eds.), *Executive function in preschool age children: Integrating measurement, neurodevelopment and translational research* (pp.299-326). Washington, DC: American Psychological Association. Retrieved from http://dx.doi. org/10.1037/14797-006

Bierman, K. L., Torres, M. M., Domitrovich, C. E., Welsh, J. A., & Gest, S. D. (2009). Behavioral and cognitive readiness for school: Cross-domain associations for children attending head start. *Social Development, 18*(2), 305-323. doi:10.1111/j.1467-9507.2008.00490.x

Bishop, D. V. M., Aamodt-Leeper, G., Creswell, C., McGurk, R., & Skuse, D. H. (2001). Individual differences in cognitive planning on the Tower of Hanoi Task: Neuropsychological maturity or measurement error? *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*(4), 551 - 556. doi:10.1017/S0021963001007247

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*(2), 647-663. doi:10.1111/j.1467-8624.2007.01019.x

Blair, C. & Dennis, T. (2010) An optimal balance: The integration of emotion and cognition in context: In S.D. Calkins, M.A. Bell (Eds.), *Child development at the intersection of emotion and cognition* (pp. 17-35). Washington, DC: American Psychological Association. doi.org/10.1037/12059-002

Blair, C. & Ursache, A. (2011). A bidirectional model of executive functions and self-regulation. In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of self-*

regulation: Research, theory, and applications (2nd ed., pp. 300-320). New York, NY: Guilford Press. Retrieved from http://www.myilibrary.com? ID=288647

Blandon, A. Y., Calkins, S. D., Grimm, K. J., Keane, S. P., & O'Brien, M. (2010). Testing a developmental cascade model of emotional and social competence and early peer acceptance. *Development and Psychopathology, 22*(4), 737-748. doi:10.1017/S0954579410000428

Bodnar, L. E., Prahme, M. C., Cutting, L. E., Denckla, M. B., & Mahone, E. M. (2007). Construct validity of parent ratings of inhibitory control. *Child Neuropsychology, 13*(4), 345-362. doi:10.1080/09297040600899867

Booth, J. N., Boyle, J. M. E., & Kelly, S. W. (2010). Do tasks make a difference? Accounting for heterogeneity of performance of children with reading difficulties on tasks of executive function: Findings from a meta-analysis. *The British Journal of Developmental Psychology,28*(1), 133-176. doi:10.1348/026151009X485432

Bowden, S. C., Fowler, K. S., Bell, R. C., Whelan, G., Clifford, C. C., Ritter, A. J., & Long, C. M. (1998). The reliability and internal validity of the Wisconsin card sorting test. *Neuropsychological Rehabilitation, 8*(3), 243-254. doi:10.1080/713755573

Brock, L. L., Rimm-Kaufman, S. E., Nathanson, L., & Grimm, K. J. (2009). The contributions of 'hot' and 'cool' executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. *Early Childhood Research Quarterly, 24*(3), 337-349. doi:10.1016/j.ecresq.2009.06.001

Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental Neuropsychology, 26*(2), 571 – 593. doi:10.1207/s15326942dn2602_3

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Cambridge, MA: Harvard University Press.

Bronfenbrenner, U. (1995). Developmental ecology through space and time: a future perspective. In P. Moen, G.H. Elder, K. Luscher, & U. Bronfenbrenner (Eds.), *Examining lives in context: Perspectives on the ecology of human development* (pp. 619-646). Washington, DC: American Psychological Association.

Bronfenbrenner, U. (2001). Bioecological theory of human development. In *International encyclopedia of the social and behavioral sciences.* (Vol. 24, pp. 6963-6970). New York, NY: Elsevier.

Bronfenbrenner, U. & Morris, P.A. (2006). The bioecological model of human development. In W. Damon & R.M. Lerner (Eds.), *Handbook of child* *psychology* (6th ed., Vol. 1, pp. 793-828). New York, NY: John Wiley. doi:10.1002/ 9780470147658.chpsy0114

Bronson, M. (2000). *Self-regulation in early childhood: Nature and nurture*. New York, NY: Guilford Press.

Brown, J. L., Roderick, T., Lantieri, L., & Aber, J. L. (2004). The Resolving Conflict Creatively Program: A school-based social and emotional learning program. In J. E. Zins, R. P. Weissberg, M. Wang, & H. J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say* (pp. 151 – 169). New York, NY: Teachers College Press.

Bryant, B.K. (1982). An index of empathy for children and adolescents. *Child Development*, *53* (2), 413 - 425. doi:10.1111/j.1467-8624.1982.tb01331.x

Brydges, C. R., Fox, A. M., Reid, C. L., & Anderson, M. (2014). The differentiation of executive functions in middle and late childhood: A longitudinal latent-variable analysis. *Intelligence, 47 (November/ December)*, 34-43. doi:10.1016/j.intell.2014.08.010

Brydges, C. R., Reid, C. L., Fox, A. M., & Anderson, M. (2012). A unitary executive function predicts intelligence in children. *Intelligence, 40*(5), 458. doi:10.1016/j.intell.2012.05.006

Buchanan, T., Heffernan, T. M., Parrott, A. C., Ling, J., Rodgers, J., & Scholey, A. B. (2010). A short self-report measure of problems with executive function suitable for administration via the Internet. *Behavior Research Methods, 42* (3), 709–714. doi:10. 3758/BRM.42.3.709

Buelow, M. T., & Barnhart, W. R. (2017). The influence of math anxiety, math performance, worry, and test anxiety on the Iowa gambling task and balloon analogue risk task. *Assessment*, *24*(1), 127-137. doi:10.1177/1073191115602554

Bujoreanu, I. S., & Willis, W. G. (2008). Developmental and neuropsychological perspectives on the Wisconsin card sorting test in children. *Developmental Neuropsychology*, *33*(5), 584-600. doi:10.1080/87565640802254364

Bull, R., Espy, K. A., & Senn, T. E. (2004). A comparison of performance on the towers of London and Hanoi in young children. *Journal of Child Psychology and Psychiatry, 45*(4), 743-754. doi:10.1111/j.1469-7610.2004.00268.x

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology, 19*(3), 273-293. doi:10.1207/S15326942DN1903_3

Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives, 8*(1), 36-41. doi:10.1111/cdep.12059

Cadima, J., Verschueren, K., Leal, T., & Guedes, C. (2016). Classroom interactions, dyadic teacher–child relationships, and self–regulation in socially disadvantaged young children. *Journal of Abnormal Child Psychology, 44*(1), 7-17. doi:10.1007/s10802-015-0060-5

Caemmerer, J. M., & Keith, T. Z. (2015). Longitudinal, reciprocal effects of social skills and achievement from kindergarten to eighth grade. *Journal of School Psychology, 53*(4), 265-281. doi:10.1016/j.jsp.2015.05.001

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*(1), 31-42. doi:10.1037/0022-0663.96.1.31

Caldarella, P., & Merrell, K. W. (1997). Common dimensions of social skills of children and adolescents: A taxonomy of positive behaviors. *School Psychology Review, 26*(2), 264–278. Retrieved from http://search.ebscohost.com.ezproxy.library.ubc.ca/login.aspx?direct=true&db=a9h&AN=9708120314&site=ehost-live&scope=site

Cameron Ponitz, C. E., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., & Morrison, F. J. (2008). Touch your toes! developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly, 23*(2), 141-158. doi:10.1016/j.ecresq.2007.01.004

Cameron Ponitz, C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology, 45*(3), 605-619. doi:10.1037/a0015365

Camerota, M., Willoughby, M.T., Kuhn, L.J, & Blair, C.B (2018). The Childhood Executive Functioning Inventory (CHEXI): Factor structure, measurement invariance, and correlates in US preschoolers. *Child Neuropsychology*, *24*(3), 322–337.

Campbell, L. E., McCabe, K. L., Melville, J. L., Strutt, P. A., & Schall, U. (2015). Social cognition dysfunction in adolescents with 22q11.2 deletion syndrome (velo-cardio-facial syndrome): Relationship with executive functioning and social competence/functioning. *Journal of Intellectual Disability Research, 59*(9), 845–859. doi:10.1111/jir.12183

Cantin, R. H., Gnaedinger, E. K., Gallaway, K. C., Hesson-McInnis, M. S., & Hund, A. M. (2016). Executive functioning predicts reading, mathematics, and theory of mind during the elementary years. *Journal of Experimental Child Psychology, 146*, 66-78. doi:10.1016/j.jecp.2016.01.014

Caprara, G. V., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P. G. (2000). Prosocial foundations of children's academic achievement. *Psychological Science, 11*(4), 302-306. doi:10.1111/1467-9280.00260

Caputi, M., Lecce, S., Pagnin, A., & Banerjee, R. (2012). Longitudinal effects of theory of mind on later peer relations: The role of prosocial behavior. *Developmental Psychology, 48*(1), 257-270. doi:http://dx.doi.org.ezproxy1.lib.asu.edu/10.1037/a0025402

Carlo, G. (2014). The development and correlates of prosocial moral behaviors. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (2nd ed., pp. 208–234). New York, NY: Psychology Press. Retrieved from http://www.myilibrary.com.ezproxy1.lib.asu.edu?ID=529223

Carlo, G., Hausmann, A., Christiansen, S., & Randall, B. A. (2003). Sociocognitive and behavioral correlates of a measure of prosocial tendencies for adolescents. *The Journal of Early Adolescence, 23*(1), 107-134. doi:10.1177/0272431602239132

Carlo, G., & Randall, B. A. (2002). The development of a measure of prosocial behaviors for late adolescents. *Journal of Youth and Adolescence, 31*(1), 31-44. doi:10.1023/A:1014033032440

Carlson, S. M. (2003). Executive function in context: Development, measurement, theory, and experience. *Monographs of the Society for Research in Child Development, 68*(3), 138-151. doi:10.1111/j.0037-976X.2003.00270.x

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science, 11*(2), 282-298. doi:10.1111/j.1467-7687.2008.00675.x

Carlson, S. M., & Wang, T. S. (2007). Inhibitory control and emotion regulation in preschool children. *Cognitive Development, 22*(4), 489-510. doi:10.1016/j.cogdev.2007.08.002

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33*(3), 386-404. doi:10.1016/0022-0965(82)90054-6

Catalano, R. F., Berglund, M. L., Jean A. M. Ryan, Lonczak, H. S., & Hawkins, J. D. (2004). Positive youth development in the United States: Research findings on evaluations of positive youth development programs. *The Annals of the American*

Academy of Political and Social Science, 591(1), 98-124. doi:10.1177/0002716203260102

Catalano, R. F., Hawkins, J. D., Berglund, M. L., Pollard, J. A., & Arthur, M. W. (2002). Prevention science and positive youth development: Competitive or cooperative frameworks? *Journal of Adolescent Health, 31*(6), 230-239. doi:10.1016/S1054-139X(02)00496-2

Catale, C., Lejeune, C., Merbah, S., & Meulemans, T. (2013). French adaptation of the childhood executive functioning inventory (CHEXI): Confirmatory factor analysis in a sample of young French-speaking Belgian children. *European Journal of Psychological Assessment, 29*(2), 149-155.

Catale, C., Meulemans, T., & Thorell, L. B. (2015). The childhood executive function inventory: Confirmatory factor analyses and cross-cultural clinical validity in a sample of 8- to 11-year-old children. *Journal of Attention Disorders, 19*(6), 489-495. doi: 1087054712470971.

Cerda, C. A., Im, M. H., & Hughes, J. N. (2014). Learning-related skills and academic achievement in academically at-risk first graders. *Journal of Applied Developmental Psychology, 35*(5), 433-443. doi:10.1016/j.appdev.2014.08.001

Chakroff, A. & Young, L. (2014). The prosocial brain. In L.M. Padilla-Walker & G. Carlo (Eds.) *Prosocial development: A multidimensional approach* (pp.90-111). Oxford Scholarship Online. doi:10.1093/acprof:oso/9780199964772.003.0005

Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. *Archives of Clinical Neuropsychology*, *21*(3), 217-227. doi:10.1016/j.acn.2005.12.002

Cheung, G.W., & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233-255.

Cicchetti, D., & Dawson, G. (2002). Multiple levels of analysis. *Development and Psychopathology, 14*(3), 417-420. doi:10.1017/S095457940200301X

Cicerone, K. D., Langenbahn, D. M., Braden, C., Malec, J. F., Kalmar, K., Fraas, M., . . . Ashman, T. (2011). Evidence-based cognitive rehabilitation: Updated review of the literature from 2003 through 2008. *Archives of Physical Medicine and Rehabilitation, 92*(4), 519-530. doi:10.1016/j.apmr.2010.11.015

Cillessen, A. H. N., Jiang, X. L., West, T. V., & Laszkowski, D. K. (2005). Predictors of dyadic friendship quality in adolescence. *International Journal of Behavioral Development, 29*(2), 165-172. doi:10.1080/01650250444000360

Cohen, J. (1988/2013). *Statistical power analysis for the behavioral sciences*. Routledge. Retrieved from http://www.myilibrary.com.ezproxy1.lib.asu.edu?ID=493929

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

Collaborative for Academic, Social, and Emotional Learning (CASEL; 2013). Effective social and emotional learning programs: Preschool and elementary school edition. Chicago, IL: CASEL. Retrieved from http://www.casel.org/preschool-and-elementary-edition-casel-guide

Common Core Standards Initiative (2012). *Common core state standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: National Governors Association. Retrieved from: http://www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf

Compas, B. E., Connor-Smith, J. K., Saltzman, H., Thomsen, A. H., & Wadsworth, M. E. (2001). Coping with stress during childhood and adolescence: Problems, progress, and potential in theory and research. *Psychological Bulletin, 127*(1), 87-127. doi:10.1037//0033-2909.127.1.87

Comrey, A.L. & Lee, H.B. (1992). *A first course in factor analysis.* Hillsdale, NJ: Lawrence Erlbaum.

Cooper, K. L. (2004) *Executive functions and social skills in children and adolescents with attention- deficit/hyperactivity disorder (AD/HD): A pilot test of Barkley's model of behavioral inhibition.* (Master's Thesis). Retrieved from ProQuest Dissertations & Theses Global (Order No. missing). Retrieved from: http://etd.fcla.edu/UF/UFE0006505/cooper_k.pdf

Côté, S., Tremblay, R. E., Nagin, D., Zoccolillo, M., & Vitaro, F. (2002). The development of impulsivity, fearfulness, and helpfulness during childhood: Patterns of consistency and change in the trajectories of boys and girls. *Journal of Child Psychology and Psychiatry, 43*(5), 609-618. doi:10.1111/1469-7610.00050

Cowan, N. (1998). Visual and auditory working memory capacity. *Trends in Cognitive Sciences, 2*(3), 77-78. doi:10.1016/S1364-6613(98)01144-9

Cragg, L., & Chevalier, N. (2012). The processes underlying flexibility in childhood. *The Quarterly Journal of Experimental Psychology, 65*(2), 209-24. doi:10.1080/17470210903204618

Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education, 3*(2), 63-68. doi:10.1016/j.tine.2013.12.001

Crowe, Beauchamp, Catroppa, & Anderson. (2011). Social function assessment tools for children and adolescents: A systematic review from 1988 to 2010. *Clinical Psychology Review, 31*(5), 767-785.doi:10.1016/j.cpr.2011.03.008

Cybele Raver, C., McCoy, D. C., Lowenstein, A. E., & Pess, R. (2013). Predicting individual differences in low-income children's executive control from early to middle childhood. *Developmental Science, 16*(3), 394-408. doi:10.1111/desc.12027

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*(11), 2037-2078. doi:10.1016/j.neuropsychologia.2006.02.006

Davidson, R.J., Dunne, J., Eccles, J.S, Engle, A., Greenberg, M., Jennings, P.,…Vago, D. (2012). Contemplative practices and mental training: Prospects for American education. *Child Development Perspectives, 6*(2), 146-153. doi:10.1111/j.1750-8606.2012.00240.x

Davis, H. L., & Pratt, C. (1996). The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology, 47*, 25–31.
de Guzman, M.R. T, Carlo, G., & Edwards, C.P. (2008). Prosocial behaviors in context: Examining the role of children's social companions. *International Journal of Behavioral Development, 32*(6), 522-530. doi:10.1177/0165025408095557

Dekker, M. C., Ziermans, T. B., Spruijt, A. M., & Swaab, H. (2017). Cognitive, parent and teacher rating measures of executive functioning: Shared and unique influences on school achievement. *Frontiers in Psychology, 8*(48), 1-13. doi: 10.3389/fpsyg.2017.00048

D'Elia, L. F., Satz, P., Uchiyama, C. L., & White, T. (1994). Color trail test: Professional manual. Odessa, FL: Psychological Assessment Resources.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System (D-KEFS)*. San Antonio, TX: The Psychological Corporation.

Denham, S. (1986). Social cognition, prosocial behavior, and emotion in preschoolers: Contextual validation. *Child Development, 57*(1), 194-201. doi: 10.2307/1130651

Derryberry, D., & Rothbart, M. K. (1988). Arousal, affect, and attention as components of temperament. *Journal of Personality and Social Psychology, 55*(6), 958-966. doi:10.1037//0022-3514.55.6.958

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of Experimental Child Psychology, 103*(2), 186-201. doi:10.1016/j.jecp.2009.01.004

de Wilde, A., Koot, H. M., & van Lier, Pol A C. (2016). Developmental links between children's working memory and their social relations with teachers and peers in the early school years. *Journal of Abnormal Child Psychology, 44*(1), 19. doi:10.1007/s10802-015-0053-4

Diamond, A. (2000). Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. *Child Development, 71*(1), 44-56. doi:10.1111/1467-8624.00117

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*(1), 135-168. doi:10.1146/annurev-psych-113011-143750

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science, 318*(5855), 1387-1388. doi:10.1126/science.1151148

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science, 333*(6045), 959-964. doi:10.1126/science.1204529

Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "Do as I say, not as I do". *Developmental Psychobiology, 29*(4), 315-334. doi: 10.1002/(SICI)1098-2302(199605)29:4<315::AID-DEV2>3.0.CO2-T

Dimond, S. J. (1980;2013). *Neuropsychology: A textbook of systems and psychological functions of the human brain*. London, England: Butterworth.

DiPerna, J.C., Lei, P., Bellinger, J., & Cheng, W. (2015). Efficacy of the Social Skills Improvement System Classwide Intervention Program (SSIS-CIP) Primary Version. *School Psychology Quarterly, 30*(1), 123-141. doi:org/10.1037/spq0000079

DiPerna, J. C., Lei, P., Cheng, W., Hart, S. C., & Bellinger, J. (2017). A cluster randomized trial of the Social Skills Improvement System-Classwide Intervention Program (SSIS-CIP) in first grade. *Journal of Educational Psychology*, doi:10.1037/edu0000191

DiStefano, C., Liu, J, Jiang, N., & Shi, D. (2018). Examination of the Weighted Root

Mean Square Residual: Evidence for Trustworthiness? *Structural Equation Modeling: A Multidisciplinary Journal, 25*(3), 453-466, DOI: 10.1080/10705511.2017.1390394

Di Tella, M., Castelli, L., Colonna, F., Fusaro, E., Torta, R., Ardito, R. B., & Adenzato, M. (2015). Theory of mind and emotional functioning in fibromyalgia syndrome: An investigation of the relationship between social cognition and executive function. *PloS One*, *10*(1), e0116542. doi:10.1371/journal.pone.0116542

Dodge, K. (1991). Emotion and social information processing. In J. Garber & K. Dodge (Eds.), *The development of emotion regulation and dysregulation*. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511663963

Dodge, K. & Garber, J. (1991). Domains of emotional regulation. In J. Garber & K. Dodge (Eds.), *The development of emotion regulation and dysregulation*. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511663963

Doyle, A. (2006). Executive functions in attention-deficit/hyperactivity disorder. *Journal of Clinical Psychiatry, 67*, 21-26.

Duckworth, A. L., & Kern, M. L. (2011). A meta-analysis of the convergent validity of self-control measures. *Journal of Research in Personality, 45*(3), 259-268. doi:10.1016/j.jrp.2011.02.004

DuPaul, G. J., & Rapport, M. D. (1991). Teacher ratings of academic skills: The development of the academic performance rating scale. *School Psychology Review*, *20*(2), 284-300.

Dusenbury, L. & Weissberg, R. P. (2016). Development and Implementation of Standards for Social and Emotional Learning in the 50 States: A Brief on Findings from CASEL's Experience. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning. Retrieved from http://www.casel.org/wp-content/uploads/2016/08/PDF-24-CASEL-Brief-Development-and-Implementation-of-SEL-Standards-in-the-50-States-April-2016.pdf

Duncan, T. E., & Duncan, S. C. (2009). The ABC's of LGM: an introductory guide to latent variable growth curve modeling. *Social and personality psychology compass*, *3*(6), 979-991.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*(1), 405-432. doi:10.1111/j.1467-8624.2010.01564.x

Dymond, R.F. (1950). Personality and empathy. *Journal of consulting psychology, 14 (5),* 343-350. Retrieved from http://login.ezproxy1.lib.asu.edu/login?url=http://search.proquest.com/docview/614266268?accountid=4485

Egeland, J., & Fallmyr, Ø. (2010). Confirmatory factor analysis of the behavior rating inventory of executive function (BRIEF): Support for a distinction between emotional and behavioral regulation. *Child Neuropsychology, 16*(4), 326-337. doi:10.1080/09297041003601462

Eilertsen, B., & Johnsen, I. M. B. (2003). Manual for Wechsler Intelligence Scale for Children–Third Edition (Norwegian version). Oslo, Norway: Assesio.

Eisenberg, N. (1986). *Altruistic emotion, cognition, and behavior*. Hillsdale, N.J: Lawrence Erlbaum Associates.

Eisenberg, N. (2012). Temperamental effortful control (self-regulation). In *Encyclopedia on early childhood development*. Retrieved from: http://www.child-encyclopedia.com/temperament/according-experts/temperamental-effortful-control-self-regulation

Eisenberg, N. & Fabes, R.A. (1991). Prosocial behavior and empathy: A multimethod developmental perspective. In M.S. Clark (Eds.), *Prosocial behavior* (pp. 34-61). Thousand Oaks, CA: Sage.

Eisenberg, N., & Fabes, R. (1998). Prosocial development. In W. Damon (Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology* (5th ed., Vol. 3, pp. 701–778). New York, NY: Wiley.

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In N. Eisenberg (Vol. Ed.) and W. Damon & R. M. Lerner (Series Eds.), *Handbook of child psychology* (6th ed., Vol.3, pp. 646–718). New York, NY: Wiley.

Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin, 101*(1), 91-119. doi:10.1037/0033-2909.101.1.91

Eisenberg, N., Smith, C.L., & Spinrad, T.L. (2011). Effortful control: Relations with emotion regulation, adjustment, and socialization in childhood. In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of self-regulation: Research, theory, and applications* (2nd ed., pp. 263-283). New York, NY: Guilford Press. Retrieved from http://www.myilibrary.com.ezproxy1.lib.asu.edu?ID=288647

Eisenberg, N., & Spinrad, T.L. (2004). Emotion-related regulation: Sharpening the definition. *Child Development, 75*(2), 334-339. doi:10.1111/j.1467-8624.2004.00674.x

Eisenberg, N., & Spinrad, T. (2014).   Multidimensionality of prosocial behavior: Rethinking the conceptualization and development of prosocial behavior. In L.M. Padilla-Walker & G. Carlo (Eds.) *Prosocial development: A multidimensional approach* (pp. 17-39). Oxford Scholarship Online. doi:10.1093/acprof:oso/ 9780199964772.003.0002

Eisenberg, N., & Spinrad, T.L., Fabes, R.A., Reiser, M., Cumberland, A., Shepard, S.A., Thompson, M. (2004). The relations of effortful control and impulsivity to children's resiliency and adjustment. *Child Development, 75*(1), 25-46. doi:10.1111/j.1467-8624.2004.00652.x

Eisenberg, N., & Spinrad, T.L., Morris, A. (2014). The development and correlates of prosocial moral behaviors. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (2nd ed., pp. 184–207). New York, NY: Psychology Press. Retrieved from http://www.myilibrary.com.ezproxy1.lib. asu.edu?ID=529223

Eisenberg, N., & Zhou, Q., (2016). Conceptions of executive functioning and regulation: When and to what degree do they overlap? In J. A. Griffin, L. S. Freund, & P. McCardle (Eds.), *Executive function in preschool age children: Integrating measurement, neurodevelopment and translational research* (pp.115-136). Washington, DC: American Psychological Association. Retrieved from http://dx.doi.org/10.1037/14797-006

Elias, M.J., Leverett, L., Duffell, J.C., Humphrey, N., Stepney, C., Ferrito, J. (2015). Integrating SEL with related development and youth approaches. In J.A. Durlak (Ed.), *Handbook of social and emotional learning: Research and practice* (pp. 33-49). New York, NY: Guilford Press.

Elias, M. J., & Moceri, D. C. (2012). Developing social and emotional aspects of learning: The American experience. *Research Papers in Education, 27*(4), 423-434. doi: 10.1080/02671522.2012.690243

Elias, M. J., Zins, J. E., Weissberg, R. P., Frey, K. S., Greenberg, M. T., Haynes, N. M., … Shriver, T.P. (1997). *Promoting social and emotional learning: Guidelines for educators.* Alexandria, VA: Association for Supervision and Curriculum Development.

Elliott, S.N., Davies, M.D., Frey, J.R., Gresham, F.M., & Cooper, G. (2017). Development and validation of a social emotional learning assessment for universal screening. *Journal of Applied Developmental Psychology*. Advance online publication. doi: 10.1016/j.appdev.2017.06.002.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*(1), 143-149. doi:10.3758/BF03203267

Eslinger, P. J. (1996). Conceptualizing, describing, and measuring components of executive function: A summary. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function* (pp. 367-395). Baltimore, MD: Brookes Publishing Co.

Evans, D. E., & Rothbart, M. K. (2007). Developing a model for adult temperament. *Journal of Research in Personality, 41*(4), 868-888. doi:10.1016/j.jrp.2006.11.002 Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience, 14*(3), 340–347. doi:10.1162/089892902317361886

Feifer, S. (2008). Integrating Response to Intervention (RTI) with neuropsychology: A scientific approach to reading. *Psychology in the Schools*, *45*(9), 812–825. https://doi.org/10.1002/pits.20328

Flanagan, D. P., & Kaufman, A. S. (2009). Essentials of WISC-IV assessment (2nd ed.). Hoboken, NJ: Wiley. Retrieved from http://www.myilibrary.com.ezproxy1.lib.asu.edu?ID=213737

Flay, B. R., Allred, C. G., & Ordway, N. (2001). Effects of the positive action program on achievement and discipline: Two matched-control comparisons. *Prevention Science, 2*(2), 71-89. doi:10.1023/A: 1011591613728

Flouri, E., & Sarmadi, Z. (2016). Prosocial behavior and childhood trajectories of internalizing and externalizing problems: The role of neighborhood and school contexts. *Developmental Psychology, 52*(2), 253-258. doi:10.1037/dev0000076

Flannery, D. J., Vazsonyi, A. T., Liau, A. K., Guo, S., Powell, K. E., Atha, H., . . . Embry, D. (2003). Initial behavior outcomes for the PeaceBuilders universal school-based violence prevention program. *Developmental Psychology, 39*(2), 292-308. doi:10.1037/0012-1649.39.2.292

Flook, L., Goldberg, S., Pinger, L., & Davidson, R. (2015). Promoting prosocial behavior and self-regulatory skills in preschool children through a mindfulness-based kindness curriculum. *Developmental Psychology, 51*(1), 44-51. doi:10.1037/a0038256

Fournet, N., Roulin, J., Monnier, C., Atzeni, T., Cosnefroy, O., Le Gall, D., & Roy, A. (2015). Multigroup confirmatory factor analysis and structural invariance with age of the behavior rating inventory of executive function (BRIEF)-French version. *Child Neuropsychology*, *21*(3), 379-20. doi:10.1080/09297049.2014.906569

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General, 133*(1), 101-135. doi:10.1037/0096-3445.133.1.101

Frijda, N. H. (1986). *The emotions*. Cambridge, UK: Cambridge University Press.
Friso-Van den Bos, I., van der Ven, S.H.G., Kroesbergen, E. H., & van Luit, J.E.H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review, 10*, 29-44. doi:10.1016/j.edurev.2013.05.003

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*(4), 483-527. doi:10.1016/0885-2014(95)90024-1

Fuhs, M. W., Nesbitt, K. T., Farran, D. C., & Dong, N. (2014). Longitudinal associations between executive functioning and academic skills across content areas. *Developmental Psychology, 50*(6), 1698-1709. doi:10.1037/a0036633

Ganesalingam, K., Yeates, K. O., Taylor, H. G., Walz, N. C., Stancin, T., & Wade, S. (2011). Executive functions and social competence in young children 6 months following traumatic brain injury. *Neuropsychology, 25*(4), 466-476. doi:10.1037/a0022768

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31-60. doi:10.1037/0033-2909.134.1.31

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3 1/2 –7 years old on a Stroop-like day-night test. *Cognition, 53*(2), 129-153. doi:10.1016/0010-0277(94)90068-X

Garcia-Barrera, M. A., Kamphaus, R. W., & Bandalos, D. (2011). Theoretical and statistical derivation of a screener for the behavioral assessment of executive functions in children. *Psychological Assessment, 23,* 64–79.

Garcia-Barrera, M., Karr, J. E., & Kamphaus, R. W. (2013). Longitudinal applications of a behavioral screener of executive functioning: Assessing factorial invariance and exploring latent growth. *Psychological Assessment, 25*(4), 1300-1313.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000a). Test Review: Behavior rating inventory of executive function. *Child Neuropsychology, 6*(3), 235-238. doi:10.1076/chin.6.3.235.3152
Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000b). BRIEF: Behavior Rating Inventory of Executive Function – Professional manual. Odessa, FL: Psychological Assessment Resources.

Gioia, G., Isquith, P., & Guy, S. (2001). Assessment of executive functions in children with neurological impairment. In R. Simeonsson & S. Rosenthal (Eds.), *Psychological and developmental assessment: Children with disabilities and chronic conditions* (pp. 317–356). New York, NY: The Guildford Press.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2015). BRIEF-2: Behavior Rating Inventory of Executive Function® – second edition. Professional manual. Lutz, FL: PAR.

Goff, P., Pratt, C. & Ong, B. (2005). The Relations Between Children's Reading Comprehension, Working Memory, Language Skills and Components of Reading Decoding in a Normal Sample. *Reading & Writing*, *18*(7), 583–616. https://doi.org/10.1007/s11145-004-7109-0

Goleman, D. P. (1995). *Emotional intelligence*. New York, NY: Bantam Books.

Goleman, D. (2004). Forward. In J.E. Zins, R.P. Weissberg, M.C. Wang & H.J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say?* (pp. vii– viii). New York, NY: Teachers College Press.

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *38*(5), 581-586. doi:10.1111/j.1469-7610.1997.tb01545.x

Gracey, F., Watson, S., McHugh, M., Swan, A., Humphrey, A., & Adlam, A. R. (2014). Age of injury, emotional problems and executive functioning in understanding disrupted social relationships following childhood acquired brain injury. *Social Care and Neurodisability,5*(3), 160-170. doi:10.1108/SCN-08-2013-0030

Graziano, P. A., Reavis, R. D., Keane, S. P., & Calkins, S. D. (2007). The role of emotion regulation in children's early academic success. *Journal of School Psychology*, *45*(1), 3-19. doi:10.1016/j.jsp.2006.09.002

Greenberg, M. T., Kusche, C. A., Riggs, N. (2004). The PATHS Curriculum: Theory and research on neuro-cognitive development and school success. In J. E. Zins, R. P. Weissberg, M. C. Wang & H. J. Walberg (Eds.). *Building academic success on social and emotional learning: What does the research say?* (pp. 170-188) New York, NY: Teachers College Press.

Gresham, F. M., & Elliott, S. N. (1990). *The Social Skills Rating System*. Circle Pines, MN: American Guidance Service [www.agsnet.com].

Gresham, F.M., & Elliott, S.N. (2008) *Social Skills Improvement System: Rating Scales.* Bloomington, MN: Pearson Assessments.

Gresham, F.M., & Elliott, S.N. (2017). *Social Skills Improvement System Social Emotional Learning Edition Rating Forms.* Minneapolis: Pearson Assessments.

Gresham, F. M., Elliott, S. N., Cook, C. R., Vance, M. J., & Kettler, R. (2010). Cross-informant agreement for ratings for social skill and problem behavior ratings: An investigation of the social skills improvement system-rating scales. *Psychological Assessment, 22*(1), 157-166. doi:10.1037/a0018124

Gresham, F. M., Elliott, S. N., & Kettler, R. J. (2010). Base rates of social skills acquisition/performance deficits, strengths, and problem behaviors: An analysis of the Social Skills Improvement System—Rating Scales. *Psychological Assessment*, *22*(4), 809-815. doi:10.1037/a0020255

Gresham, F. M., Elliott, S. N., Metallo, S., Byrd, S., Wilson, E., Erickson, M., & Altman, R.A. (in press). *Comparability of SSIS Rating Scales to the SSIS Social Emotional Learning Rating Forms: Content and psychometric comparisons across elementary and secondary levels.* Manuscript submitted for publication.

Gresham, F. M., Elliott, S. N., Vance, M. J., & Cook, C. R. (2011). Comparability of the social skills rating system to the social skills improvement system: Content and psychometric comparisons across elementary and secondary age levels. *School Psychology Quarterly, 26*(1), 27-44. doi:10.1037/a0022662

Gresham, F., Metallo, S., Byrd, S., Wilson, E., Erickson, M., Cassidy, K., . . . Altman, R. (2018). Psychometric fundamentals of the Social Skills Improvement System: Social–Emotional Learning Edition Rating Forms. *Assessment for Effective Intervention*. doi: 10.1177/1534508418808598

Grimm, K. J., Ram, N., & Estabrook, R. (2017). Growth modeling: Structural equation and multilevel modeling approaches. New York, NY: Guilford Publications. Retrieved from: http://www.myilibrary.com.ezproxy1.lib.asu.edu?ID=962687

Hackman, D. A., Gallop, R., Evans, G. W., & Farah, M. J. (2015). Socioeconomic status and executive function: Developmental trajectories and mediation. *Developmental Science, 18*(5), 686-702. doi:10.1111/desc.12246

Henk, C. M., & Castro-Schilo, L. (2016). Preliminary detection of relations among dynamic processes with two-occasion data. *Structural Equation Modeling: A Multidisciplinary Journal, 23*(2), 180-193. doi:10.1080/10705511.2015.1030022

Hartup, W.W. (1985). Relationships and their significance in cognitive development. In R.A. Hinde, A.N. Perret-Clermont, & J. Stevenson-Hinde (Eds.) *Social relationships and cognitive development* (pp.66-82). New York, NY: Oxford University Press.

Heatly, M. C., Bachman, H. J., & Votruba-Drzal, E. (2015). Developmental patterns in the associations between instructional practices and children's math trajectories in elementary school. *Journal of Applied Developmental Psychology, 41*, 46-59. doi:10.1016/j.appdev.2015.06.002

Heaton, R. (1981). Wisconsin Card Sorting Test (WCST). Odessa, FL: Professional Assessment Resources.

Hensler, M., Wolfe, K., Lebensburger, J., Nieman, J., Barnes, M., Nolan, W., . . . Madan-Swain, A. (2014). Social skills and executive function among youth with sickle cell disease: A preliminary investigation. Journal of Pediatric Psychology, 39(5), 493-500. doi:10.1093/jpepsy/jst138

Heverly-Fitt, S., Rubin, K. H., Dennis, M., Taylor, H. G., Stancin, T., Gerhardt, C.A., …Yeates, K. O. (2016). Investigating a proposed model of social competence in children with traumatic brain injuries. *Journal of Pediatric Psychology, 41*(2), 235-243. doi:10.1093/jpepsy/jsv085

Hodges, S.D., & Myers, M.W. (2007). Empathy. In R.F. Baumeister & K.D. Vohs (Eds.), *Encyclopedia of social psychology* (pp. 297-299). Thousand Oaks, CA: Sage Publications. doi:10.4135/9781412956253

Hoefling, L. K. (2016). *Exploring the construct validity of the behavioral rating inventory of executive function (BRIEF): Comparing variance associated with common method versus similar constructs* (Doctoral dissertation). Available from ProQuest Dissertations & Theses Global (Order No. 10106111). Retrieved from http://login.ezproxy1.lib.asu.edu/login?url=http://search.proquest.com. ezproxy1.lib.asu.edu/docview/1791122942?accountid=4485

Hoffman, D. M. (2009). Reflecting on social emotional learning: A critical perspective on trends in the United States. *Review of Educational Research, 79*(2), 533-556. doi:10.3102/0034654308325184

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences, 16*(3), 174-180. doi:10.1016/j.tics.2012.01.006

Homack, S., & Riccio, C. A. (2004). A meta-analysis of the sensitivity and specificity of the stroop color and word test with children. *Archives of Clinical Neuropsychology, 19*(6), 725-743. doi:10.1016/j.acn.2003.09.003

Hughes, C., Ensor, R., Wilson, A., & Graham, A. (2010). Tracking executive function across the transition to school: A latent variable approach. *Developmental Neuropsychology, 35*(1), 20-36. doi:10.1080/87565640903325691

Hughes, C., & Devine, R. T. (2015). Individual differences in theory of mind from preschool to adolescence: Achievements and directions. *Child Development Perspectives, 9*(3), 149-153. doi:10.1111/cdep.12124

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017-2036. doi:10.1016/j.neuropsychologia.2006.01.010

Isquith, P.K., Gioia, G.A., & PARi staff (2013). Behavior rating inventory of executive function®: Sample teacher form interpretive report. Lutz, FL: Pari Connect. Retrieved from http://www4.parinc.com/WebUploads/samplerpts/BRIEF_Teacher_Interp_PiC.pdf

Iyer, R. V., Kochenderfer-Ladd, B., Eisenberg, N., & Thompson, M. (2010). Peer victimization and effortful control: Relations to school engagement and academic achievement. *Merrill-Palmer Quarterly, 56*(3), 361-387. doi:10.1353/mpq.0.0058

Jacob, R., & Parkinson, J. (2015). The potential for school-based interventions that target executive function to improve academic achievement: A review. *Review of Educational Research, 85*(4), 512-552. doi:10.3102/0034654314561338

Jiménez, A., & Lucas-Molina, B. (2019). Dimensional structure and measurement invariance of the BRIEF-2 across gender in a socially vulnerable sample of primary school-aged children. *Child Neuropsychology*, *25*(5), 636–647.

Jones, S. M., Brown, J. L., & Lawrence Aber, J. (2011). Two-Year impacts of a universal School-Based Social-Emotional and literacy intervention: An experiment in translational developmental research. *Child Development*, *82*(2), 533-554. doi:10.1111/j.1467-8624.2010.01560.x

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology, 45*(3), 850-867. doi:10.1037/a0014939

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review, 17*(3), 213-233. doi:10.1007/s11065-007-9040-z

Kaller, C. P., Rahm, B., Spreer, J., Mader, I., & Unterrainer, J. M. (2008). Thinking around the corner: The development of planning abilities. *Brain and Cognition, 67*(3), 360-370. doi:10.1016/j.bandc.2008.02.003

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*(2), 189-217. doi:http://dx.doi.org.ezproxy1.lib.asu.edu/10.1037/0096-3445.133.2.189

Kerr, A., & Zelazo, P. D. (2004). Development of "hot" executive function: The children's gambling task. *Brain and Cognition, 55*(1), 148-157. doi:10.1016/S0278-2626(03)00275-6

Kettler, R. J., Elliott, S. N., Beddow, P. A., Compton, E., McGrath, D., Kaase, K. J., . . . Hinton, K. (2010). What do alternate assessments of alternate academic achievement standards measure? A Multitrait—Multimethod analysis. *Exceptional Children, 76*(4), 457-474. doi:10.1177/001440 291007600405

Kiley-Brabeck, K., & Sobin, C. (2006). Social skills and executive function deficits in children with the 22q11deletion syndrome. *Applied Neuropsychology, 13*(4), 258-268. doi:10.1207/s15324826an1304_7

Kimberg, D. Y., D'Esposito, M., & Farah, M. J. (1997). Cognitive functions in the prefrontal cortex: Working memory and executive control. *Current Directions in Psychological Science, 6*(6), 185-192. doi:10.1111/1467-8721.ep10772959

Kimberg, D. Y., & Farah, M. J. (1993). A unified account of cognitive impairments following frontal lobe damage: The role of working memory in complex, organized behavior. *Journal of Experimental Psychology: General, 122*(4), 411-428. doi:10.1037/0096-3445.122.4.411

Kingdon, D., Serbin, L. A., & Stack, D. M. (2017;2016;). Understanding the gender gap in school performance among low-income children: A developmental trajectory analysis. *International Journal of Behavioral Development, 41*(2), 265-274. doi:10.1177/0165025416631836

Knafo, A., & Plomin, R. (2006). Prosocial behavior from early to middle childhood: Genetic and environmental influences on stability and change. *Developmental Psychology, 42*(5), 771-786. doi:10.1073/0012-1649.42.5.771

Kokko, K., Tremblay, R. E., Lacourse, E., Nagin, D. S., & Vitaro, F. (2006). Trajectories of prosocial behavior and physical aggression in middle childhood: Links to adolescent school dropout and physical violence. *Journal of Research on Adolescence, 16*(3), 403-428. doi:10.1111/j.1532-7795.2006.00500.x

Korkman, M., Kirk, U., & Kemp, S. (2007). NEPSY Second Edition (NEPSY-II) . San Antonio, TX:Harcourt Assessment.

Kramer, A. F., Larish, J. F., & Strayer, D. L. (1995). Training for attentional control in dual task settings: A comparison of young and old adults. *Journal of Experimental Psychology: Applied, 1*(1), 50-76. doi:10.1037/1076-898X.1.1.50

Kroesbergen, E. H., van 't Noordende, Jaccoline E, & Kolkman, M. E. (2014). Training working memory in kindergarten children: Effects on working memory and early numeracy. *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence, 20*(1), 23-37. doi:10.1080/09297049.2012.736483

Ladd, G. W. (2005). *Children's peer relationships and social competence: A century of progress.* New Haven: CT: Yale University Press.

Ladd, G. W., Kochenderfer-Ladd, B., Visconti, K. J., Ettekal, I., Sechler, C. M., & Cortes, K. I. (2014). Grade-school children's social collaborative skills: Links with partner preference and achievement. *American Educational Research Journal, 51*(1), 152-183. doi:10.3102/0002831213507327

Ladd, G. W., & Profilet, S. M. (1996). The child behavior scale: A teacher-report measure of young children's aggressive, withdrawn, and prosocial behaviors. *Developmental Psychology, 32*(6), 1008-1024. doi:10.1037/0012-1649.32.6.1008

Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology, 108*(3), 677-692. doi:10.1016/j.jecp.2010.11.001

Larson, G. E., Merritt, C. R., & Williams, S. E. (1988). Information processing and intelligence: Some implications of task complexity. *Intelligence, 12*(2), 131-147. doi:10.1016/0160-2896(88)90012-8

LeBuffe, P.A., Naglieri, J.A., & Shapiro, V.B. (2012). *The Devereux Student Strengths Assessment – Second Step Edition (DESSA-SSE) Assessment, Technical Manual, and User's Guide*. Charlotte, NC: Apperson, Inc.

Lee, D. S. (2016). *A longitudinal analysis of executive functions, learning-related skills, and mathematics achievement* (Doctoral dissertation). Available from ProQuest Dissertations & Theses Global (Order No. 10168591). Retrieved from http://login.ezproxy1.lib.asu.edu/login?url=http://search. proquest.com.ezproxy1.lib.asu.edu/docview/1832353374?accountid=4485

Lee, J. (2010). Tripartite growth trajectories of reading and math achievement: Tracking national academic progress at primary, middle, and high school levels. *American Educational Research Journal, 47*(4), 800-832. doi:10.3102/0002831210365009

Lee, K., Ng, S. F., Pe, M. L., Ang, S. Y., Hasshim, M.N.A.N., & Bull, R. (2012). The cognitive underpinnings of emerging mathematical skills: Executive functioning, patterns, numeracy, and arithmetic: Cognitive underpinnings. *British Journal of Educational Psychology, 82*(1), 82-99. doi:10.1111/j.2044-8279.2010.02016.x

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology, 21*(1), 59-80. doi:10.1348/026151003321164627

Lemaire, P., & Lecacheur, M. (2011). Age-related changes in children's executive functions and strategy selection: A study in computational estimation. *Cognitive Development, 26*(3), 282-294. doi:10.1016/j.cogdev.2011.01.002

Lemberger, M. E., Selig, J. P., Bowers, H., & Rogers, J. E. (2015). Effects of the student success skills program on executive functioning skills, feelings of connectedness, and academic achievement in a predominantly Hispanic, Low-Income middle school district. *Journal of Counseling & Development, 93*(1), 25-37. doi:10.1002/j.1556-6676.2015.00178.x

Lennon, R., & Eisenberg, N. (1987). Gender and age differences in empathy and sympathy. In N. Eisenberg & J. Strayer (Eds.), Empathy and its development (pp. 195–217). New York: Cambridge University Press.

Lezak, M.D. (2004). *Neuropsychological assessment* (4th ed.). New York, NY: Oxford University Press.

Li, C. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936-949.

Liew, J. (2012). Effortful control, executive functions, and education: Bringing self-regulatory and social- emotional competencies to the table. *Child Development Perspectives, 6*(2), 105-111. doi:10.1111/j.1750-8606.2011.00196.x

Li-Grining, C. P., Votruba-Drzal, E., Maldonado-Carreño, C., & Haas, K. (2010). Children's early approaches to learning and academic trajectories through fifth grade. *Developmental Psychology, 46*(5), 1062-1077. doi:10.1037/a0020066

Little, M. (2017). Racial and socioeconomic gaps in executive function skills in early elementary school: Nationally representative evidence from the ECLS-K:2011. *Educational Researcher, 46*(2), 103- 109. doi:10.3102/0013189X17698700

Little, T. (2013). *Longitudinal structural equation modeling.* New York, NY: Guilford Press.

Liu, Y., Millsap, R. E., West, S. G., Tein, J. Y., Tanaka, R., & Grimm, K. J. (2017). Testing measurement invariance in longitudinal data with ordered-categorical measures. *Psychological methods, 22*(3), 486-506.

Llorente, A. M., Williams, J., Satz, P., & D'Elia, L. F. (2003). Children's colour trail test. Lutz: Psychological Assessment Resources, Inc.

Lockwood PL, Seara-Cardoso A, Viding, E. (2014). Emotion regulation moderates the association between empathy and prosocial behavior. *PLoSONE 9*(5): e96555. doi:10.1371/journal.pone.0096555

Logan, G. D. (1994). On the ability to inhibit thought and action: A users guide to the stop-signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189–239). San Diego, CA: Academic Press.

Low, S., Cook, C., Smolkowski, K., & Buntain-Ricklefs, J. (2015). Promoting social-emotional competence: An evaluation of the elementary version of Second Step®. *Journal of School Psychology, 53*(6), 463-477. doi:10.1016/j.jsp.2015.09.002

Low, S., Smolkowski, K., Cook, C., & Desfosses, D. (2019). Two-year impact of a universal social-emotional learning curriculum: Group differences from developmentally sensitive trends over time. Developmental Psychology, 55(2), 415-433. doi:http://dx.doi.org.ezproxy1.lib.asu.edu/10.1037/dev0000621

Luria, A.R. (1966). *Higher cortical functions in man*. New York: Basic Books. Macdonald, J. A., Beauchamp, M. H., Crigan, J. A., & Anderson, P. J. (2014). Age-related differences in inhibitory control in the early school years. *Child Neuropsychology*, *20*(5/6), 509-526. doi:10.1080/09297049.2013.822060

Lyons Usher, A. M., Leon, S. C., Stanford, L. D., Holmbeck, G. N., & Bryant, F. B. (2016). Confirmatory factor analysis of the Behavior Rating Inventory of Executive Functioning (BRIEF) in children and adolescents with ADHD. *Child Neuropsychology*, *22*(8), 907–918.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163-203. doi:10.1037/0033-2909.109.2.163

MacLeod, C. M. (1992). The stroop task: The "gold standard" of attentional measures. *Journal of Experimental Psychology: General, 121*(1), 12-14. doi:10.1037/0096-3445.121.1.12

Mahone, E. M., Hagelthorn, K. M., Cutting, L. E., Schuerholz, L. J., Pelletier, S. F., Rawlins, C., & Denckla, M. B. (2002). Effects of IQ on executive function measures in children with ADHD. *Child Neuropsychology*, *8*(1), 52-65. doi:10.1076/chin.8.1.52.8719

Malecki, C. K., & Elliott, S. N. (2002). Children's social behaviors as predictors of academic achievement: A longitudinal analysis. *School Psychology Quarterly, 17*(1), 1-23. doi:10.1521/scpq.17.1.1.19902

Männamaa, M., Kikas, E., Peets, K., & Palu, A. (2012). Cognitive correlates of math skills in third-grade students. *Educational Psychology*, *32*(1), 21-44. doi:10.1080/01443410.2011.621713

Maricle, D.E. & Avirett, E. (2012). The role of cognitive and intelligence tests in the assessment of executive functions. In D.P. Flanagan & P.L. Harrison (Eds). *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 820-838). New York, NY: Guilford Press.

Masten, A. S., & Cicchetti, D. (2010). Developmental cascades. *Development and Psychopathology, 22*(3), 491-495. doi:10.1017/S0954579410000222

Masten, A. S., Herbers, J. E., Desjardins, C. D., Cutuli, J. J., McCormick, C. M., Sapienza, J. K., . . . Zelazo, P. D. (2012). Executive function skills and school success in young children experiencing homelessness. *Educational Researcher, 41*(9), 375-384. doi:10.3102/0013189X12459883

Matthews, J. S., Kizzie, K. T., Rowley, S. J., & Cortina, K. (2010). African Americans and boys: Understanding the literacy gap, tracing academic trajectories, and evaluating the role of learning-related skills. *Journal of Educational Psychology, 102*(3), 757-771. doi:10.1037/a0019616

McAuley, T., Chen, S., Goos, L., Schachar, R., & Crosbie, J. (2010). Is the behavior rating inventory of executive function more strongly associated with measures of impairment or executive function? *Journal of the International Neuropsychological Society, 16*(3), 495-505. doi:10.1017/S1355617710000093

McCann, S. J. (2010). *Latent growth curve analyses of the development of executive functions in a prospective, longitudinal cohort of children with prenatal cocaine exposure* (Doctoral dissertation). Available from ProQuest Dissertations & Theses Global. (Order No. 3480463). Retrieved from http://ezproxy.library.ubc.ca/login?url=https://search-proquest com.ezproxy.library. ubc.ca/docview/899273899?accountid=14656

McClelland, M. M., Acock, A. C., & Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school.

Early Childhood Research Quarterly, 21(4), 471-490. doi:10.1016/j.ecresq.2006.09.003

McClelland, M. M., & Cameron, C. E. (2012). Self-Regulation in early childhood: Improving conceptual clarity and developing ecologically valid measures. *Child Development Perspectives, 6*(2), 136-142. doi:10.1111/j.1750-8606.2011.00191.x

McClelland, M. M., Cameron, C. E., Wanless, S. B., & Murray, A. (2007). Executive function, self- regulation, and social-emotional competence: Links to school readiness. In O. N. Saracho & B. Spodek (Eds.), *Contemporary perspectives on research in social learning in early childhood education*, (pp. 83-107). Charlotte, NC: Information Age.

McClelland, M. M., & Morrison, F. J. (2003). The emergence of learning-related social skills in preschool children. *Early Childhood Research Quarterly, 18*(2), 206-224. doi:10.1016/S0885-2006(03)00026-7

McFarland, J., Hussar, B., Zhang, J., Wang, X., Wang, K., Hein, S., Diliberti, M., Forrest Cataldi, E., Bullock Mann, F., and Barmer, A. (2019). *The Condition of Education 2019* (NCES 2019-144). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved [date] from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2019144.

McNeil, N. M. (2007). U-shaped development in math: 7-year-olds outperform 9-year-olds on equivalence problems. *Developmental Psychology, 43*(3), 687-695. doi:10.1037/0012-1649.43.3.687

Meuwese, R., Cillessen, A. H. N., & Güroğlu, B. (2016). Friends in high places: A dyadic perspective on peer status as predictor of friendship quality and the mediating role of empathy and prosocial behavior. *Social Development*. Advance online publication. doi:10.1111/sode.12213

Miles, S. B., & Stipek, D. (2006). Contemporaneous and longitudinal associations between social behavior and literacy achievement in a sample of low-income elementary school children. *Child Development, 77*(1), 103-117. doi:10.1111/j.1467-8624.2006.00859.x

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167-202. doi:10.1146/annurev.neuro.24.1.167

Miller, M., Loya, F., & Hinshaw, S. P. (2013). Executive functions in girls with and without childhood ADHD: Developmental trajectories and associations with symptom change. *Journal of Child Psychology and Psychiatry, 54*(9), 1005-1015. doi:10.1111/jcpp.12074

Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin, 27*(3), 272-277. doi:10.1093/oxfordjournals.bmb.a070866

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100. doi:10.1006/cogp.1999.0734

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*(1), 8-14. doi:10.1177/0963721411429458

Monette, S., Bigras, M., & Lafrenière, M. (2015). Structure of executive functions in typically developing kindergarteners. *Journal of Experimental Child Psychology, 140*, 120-139. doi:10.1016/j.jecp.2015.07.005

Monette, S., Bigras, M., & Guay, M. (2011). The role of the executive functions in school achievement at the end of grade 1. *Journal of Experimental Child Psychology, 109*(2), 158-173. doi:10.1016/j.jecp.2011.01.008

Monsell, S. (1996). Control of mental processes. In V. Bruce (Ed.), *Unsolved mysteries of the mind: Tutorial essays in cognition* (pp.93-148). Hove, East Sussex, UK: Erlbaum (UK) Taylor & Francis.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*(3), 134-140. doi:10.1016/S1364- 6613(03)00028-7

Moran, L. M., Bigler, E., Dennis, M., Gerhardt, C. A., Rubin, K. H., Stancin, T., . . .

Moriguchi, Y. (2014). The early development of executive function and its relation to social interaction: A brief review. *Frontiers in Psychology, 5*(388), 1-6. doi:10.3389/fpsyg.2014.00388

Morris, R.D. (1996). Relationships and distinctions among the concepts of attention, memory, and executive function: A developmental perspective. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function* (pp. 11-16). Baltimore, MD: Brookes Publishing Co.

Morrison, F. J., & Grammar, J.K. (2016). Conceptual clutter and measurement mayhem: A proposal for across disciplinary approach to conceptualizing and measuring executive function. In J. A. Griffin, L. S. Freund, & P. McCardle (Eds.), *Executive function in preschool age children: Integrating measurement,*

neurodevelopment and translational research (pp.327-348). Washington, DC: American Psychological Association.
Retrieved from http://dx.doi. org/10.1037/14797-006

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? the promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review, 18*(1), 46-60. doi:10.3758/s13423-010-0034-0

Morrison, F. J., Ponitz, C.C., McClelland, M.M (2010). Self-regulation and academic achievement in the transition to school. In S.D. Calkins & M.A. Bell (2010). *Child development at the intersection of emotion and cognition* (pp.203-224). Washington, DC: American Psychological Association. doi:10.1037/12059-000

Morton, J.B., & Carlson, S.M. (2017). The bilingual advantage: evidence and alternative views (pp.111-117). In M. Hoskyn, G. Iarocci, & A.R. Young (Eds.) *Executive functions in children's everyday lives: A handbook for professionals in applied psychology*. New York, NY: Oxford University Press. doi:10.1093/acprof:oso/9780199980864.001.0001

Mungas, D., Widaman, K., Zelazo, P. D., Tulsky, D., Heaton, R. K., Slotkin, J., & …Gershon, R. C. (2013). VII. NIH toolbox cognition battery (CB): Factor structure for 3 to 15 year olds. *Monographs of the Society for Research in Child Development, 78*(4), 103-118. doi:10.1111/mono.12037

Muthén, L.K. & Muthén, B.O. (1998-2017). Mplus User's Guide. Eigth Edition. Los Angeles, CA: Muthén & Muthén

Nantel-Vivier, A., Pihl, R. O., Côté, S., & Tremblay, R. E. (2014). Developmental association of prosocial behaviour with aggression, anxiety and depression from infancy to preadolescence. *Journal of Child Psychology and Psychiatry, 55* (10), 1135–1144. http://dx.doi.org/10.1111/jcpp.12235

Nesbitt, K. T., Farran, D. C., & Fuhs, M. W. (2015). Executive function skills and academic achievement gains in prekindergarten: Contributions of learning-related behaviors. *Developmental Psychology, 51*(7), 865-878. doi:http://dx.doi.org.ezproxy1.lib.asu.edu/10.1037/dev0000021

Nigg, J. T., Quamma, J. P., Greenberg, M. T., & Kusche, C. A. (1999). A two-year longitudinal study of neuropsychological and cognitive performance in relation to behavioral problems and competencies in elementary school children. Journal of Abnormal Child Psychology, 27(1), 51-63. doi:10.1023/A:1022614407893

Noble, K. G., Norman, M. F., & Farah, M. J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Developmental Science, 8*(1), 74-87. doi:10.1111/j.1467-7687.2005.00394.x

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. 4, pp. 1–18). New York, NY: Plenum Press. doi:10.1007/978-1-4757-0629-1

Neuenschwander, R., Röthlisberger, M., Cimeli, P., & Roebers, C. M. (2012). How do different aspects of self-regulation predict successful adaptation to school? *Journal of Experimental Child Psychology, 113*(3), 353-371. doi:10.1016/j.jecp.2012.07.004

Normand, S., & Tannock, R. (2014). Screening for working memory deficits in the classroom: The psychometric properties of the working memory rating scale in a longitudinal school-based study. *Journal of Attention Disorders, 18*(4), 294-304. doi:10.1177/1087054712445062

Padilla-Walker, L.M. & Carlo, G. (2014). The study of prosocial behavior: Past, present, and future. In L.M. Padilla-Walker & G. Carlo (Eds.) *Prosocial development: A multidimensional approach* (pp. 3-24). Oxford Scholarship Online. doi: 10.1093/acprof:oso/9780199964772.003.0001

Payton, J. W., Wardlaw, D. M., Graczyk, P. A., Bloodworth, M. R., Tompsett, C. J., & Weissberg, R. P. (2000). Social and emotional learning: A framework for promoting mental health and reducing risk behavior in children and youth. *Journal of School Health, 70*(5), 179-185. doi:10.1111/j.1746-1561.2000.tb06468.x

Payton, J., Weissberg, R. P., Durlak, J. A., Dymnicki, A. B., Taylor, R. D., Schellinger, K. B., & Pachan, M. (2008). *The positive impact of social and emotional learning for kindergarten to eighth-grade students: Findings from three scientific reviews.* Chicago, IL: Collaborative for Academic, Social, and Emotional Learning. Retrieved from http://www.casel.org/the-positive-impact-of-social -and-emotional-learning-for-kindergarten-to-eighth-grade-students-findings-from-three-scientific-reviews/

Pellegrino, J. W., & Hilton, M. L. (Eds.). (2012). Education for life and work: Developing transferable knowledge and skills in the 21st Century. Washington, DC: The National Academies Press. Retrieved from http://ebookcentral.proquest.com.ezproxy1.lib.asu.edu/lib/asulib-ebooks/detail.action?docID=3379216

Penner, L. A., Dovidio, J. F., Piliavin, J. A., & Schroeder, D. A. (2005). Prosocial behavior: Multilevel perspectives. *Annual Review of Psychology, 56*(1), 365-392. doi:10.1146/annurev.psych.56.091103.070141

Peters, C., Algina, J., Smith, S. W., & Daunic, A. P. (2012). Factorial validity of the behavior rating inventory of executive function (BRIEF)-teacher form. *Child Neuropsychology, 18*(2), 168-181. doi:10.1080/09297049.2011.594427

Peterson, E. & Welsh, M.C. (2014). The development of hot and cool executive functions in childhood and adolescence: Are we getting warmer? In S. Goldstein & J.A. Naglieri (Eds.) *Handbook of executive functioning* (pp. 45-66). New York, NY: Springer. doi:10.1007/978-1-4614-8106-5

Piaget, J. (1976). The grasp of consciousness: Action and concept in the young child (S. Wedgewood, Trans.). Cambridge, MA: Harvard University Press (Original work published 1974).

Poorthuis, A., Thomaes, S., Denissen, J. J. A., van Aken, M. A. G., & Orobio de Castro, B. (2012). Prosocial tendencies predict friendship quality, but not for popular children. *Journal of Experimental Child Psychology, 112*(4), 378-388. doi:10.1016/j.jecp.2012.04.002

Posner, M. I., & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology, 12*(3), 427-441. doi:10.1017/S0954579400003096

Poulin-Dubois, D., Blaye, A., Coutya, J., & Bialystok, E. (2011). The effects of bilingualism on toddlers' executive functioning. *Journal of Experimental Child Psychology, 108*(3), 567-579. doi:10.1016/j.jecp.2010.10.009

Putnick, D.L., & Bornstein, M.H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. Developmental Review, 41, 71-90. DOI: 10.1016/j.dr.2016.06.004

Qian, Y., Shuai, L., Chan, R. C. K., Qian, Q., & Wang, Y. (2013). The developmental trajectories of executive function of children and adolescents with attention deficit hyperactivity disorder. *Research in Developmental Disabilities, 34*(5), 1434-1445. doi:10.1016/j.ridd.2013.01.033

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavioral Assessment System for Children—Second Edition*. Minneapolis, MN: Pearson Assessments.

Reynolds, C.R., & MacNeill Horton Jr., A. (2014). The neuropsychology of executive functioning and the DSM-5. In S. Goldstein & J.A. Naglieri (Eds.) *Handbook of executive functioning* (pp.88-106). New York, NY: Springer. doi:10.1007/978-1-4614-8106-5

Rhoades, B. L., Warren, H. K., Domitrovich, C. E., & Greenberg, M. T. (2011). Examining the link between preschool social–emotional competence and first grade academic achievement: The role of attention skills. *Early Childhood Research Quarterly, 26*(2), 182-191. doi:10.1016/j.ecresq.2010.07.003

Richaud, M. C., Mesurado, B. K., & Cortada, A. (2012). Analysis of dimensions of prosocial behavior in an Argentinean sample of children. *Psychological Reports, 111*(3), 687-696. doi:10.2466/10.11.17.PR0.111.6.687-696

Ridderinkhof, K. R., & van der Molen, M. W. (1995). A psychophysiological analysis of developmental differences in the ability to resist interference. *Child Development, 66*(4), 1040-1056. doi:10.2307/1131797

Riggs, N., Greenberg, R., Kusché, M., & Pentz, T. (2006). The Mediational Role of Neurocognition in the Behavioral Outcomes of a Social-Emotional Prevention Program in Elementary School Students: Effects of the PATHS Curriculum. *Prevention Science, 7*(1), 91-102.

Riggs, N. R., Jahromi, L. B., Razza, R. P., Dillworth-Bart, J. E., & Mueller, U. (2006). Executive function and the promotion of social–emotional competence. *Journal of Applied Developmental Psychology, 27*(4), 300-309. doi:10.1016/j.appdev.2006.04.002

Rimm-Kaufer, S.E. & Hulleman, C.S. (2015). SEL in elementary school settings: Identifying mechanisms that matter. In J.A. Durlak (Ed.) *Handbook of social and emotional learning: Research and practice* (pp. 151-166). New York, NY: The Guilford Press. Retrieved from http://search.ebscohost.com.spx?direct=true&db=nlebk&AN=969923&site=ehost-live&scope=site

Rimm-Kaufman, S. E., Fan, X., Chiu, Y., & You, W. (2007). The contribution of the responsive classroom approach on children's academic achievement: Results from a three-year longitudinal study. *Journal of School Psychology, 45*(4), 401-421. doi:10.1016/j.jsp.2006.10.003

Roberts, R., & Pennington, B. (1996). An interactive framework for examining prefrontal cognitive processes. *Developmental Neuropsychology, 12*(1), 105-126. doi:10.1080/87565649609540642

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207-231. doi:10.1037/0096-3445.124.2.207

Romine, C. B., Lee, D., Wolfe, M. E., Homack, S., George, C., & Riccio, C. A. (2004). Wisconsin card sorting test with children: A meta-analytic study of sensitivity and specificity. *Archives of Clinical Neuropsychology, 19*(8), 1027-1041. doi:10.1016/j.acn.2003.12.009

Romano, E., Babchishin, L., Pagani, L. S., & Kohen, D. (2010). School readiness and later achievement: Replication and extension using a nationwide Canadian survey. *Developmental Psychology, 46*(5), 995-1007. doi:10.1037/a0018880

Rose-Krasnor, L. (1997). The nature of social competence: A theoretical review. *Social Development, 6*(1), 111-135. doi:10.1111/1467-9507.00029

Rose-Krasnor, L., & Denham, S. (2009). Social-emotional competence in early childhood. In K. H. Rubin, W.M. Bukowski, & B Laursen (Eds.), *Handbook of peer interactions, relationships, and groups* (pp.162-179). Guilford Press: New York, NY. Retrieved from http://ezproxy.library.ubc.ca/login?url=http://search.proquest.com.ezproxy.library.ubc.ca/docview/199766951?accountid=14656

Roth, R. M., Erdodi, L. A., McCulloch, L. J., & Isquith, P. K. (2015). Much ado about norming: The behavior rating inventory of executive function. *Child Neuropsychology, 21*(2), 225-233. doi:10.1080/09297049.2014.897318

Rothbart, M. K. (2011). *Becoming who we are: Temperament and personality in development*. New York, NY: Guilford Press. Retrieved from http://www.myilibrary.com.ezproxy1.lib.asu.edu?ID=302097

Rothbart, M. K., Ahadi, S. A., & Evans, D. E. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality and Social Psychology, 78*(1), 122-135. doi:10.1037//0022-3514.78.1.122

Rothbart, M., Ahadi, S., & Hershey, K. (1994). Temperament and social behavior in childhood. *Merrill-Palmer Quarterly, 40*(1), 21–39. Retrieved from http://www.jstor.org.ezproxy1.lib.asu.edu/stable/23087906

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The children's behavior questionnaire. *Child Development, 72*(5), 1394-1408. doi:10.1111/1467-8624.00355

Rothbart, M.K. & Bates, J.E. (2006). Temperament. In W. Damon (Editor in Chief) & N. Eisenberg (Vol. Ed.), *Handbook of Child Psychology* (6th ed., Vol.3, p.99). Hoboken, N.J: John Wiley & Sons. doi:10.1002/9780470147658.chpsy0303

Rothbart, M. K.; Bates, J. E., & Eisenberg, N. (1998). Temperament. In W. Damon (Ed.) *Handbook of Child Psychology* (5th ed., Vol.3, pp. 105-176). New York, NY: J. Wiley.

Rothbart, M. K., & Rueda, M. R. (2005). The development of effortful control. In U. Mayr, E. Awh, & S. W. Keele (Eds.), *Developing individuality in the human brain: A festschrift honoring Michael I. Posner - May 2003 (*pp. 167-188). Washington, DC: American Psychological Association Press.

Röthlisberger, M., Neuenschwander, R., Cimeli, P., Michel, E., & Roebers, C. M. (2012). Improving executive functions in 5- and 6-year-olds: Evaluation of a small group intervention in prekindergarten and kindergarten children. *Infant and Child Development, 21*(4), 411-429.

Saarni, C. (1984). An observational study of children's attempts to monitor their expressive behavior. *Child Development, 55*(4), 1504-1513. doi:10.2307/1130020

Schmid, C., Zoelch, C., & Roebers, C. M. (2008). Das Arbeitsgedächtnis von 4- bis 5-jährigen Kindern: Theoretische und empirische Analyse seiner Funktionen [Working memory in 4- to 5-year-olds: Theoretical and empirical analyses of its functions]. Zeitschrift für Entwicklungspsychologie und Pädagogische *Psychologie, 1,* 2–12.

Scourfield, J., John, B., Martin, N., & McGuffin, P. (2004). The development of prosocial behaviour in children and adolescents: A twin study. *Journal of Child Psychology and Psychiatry, 45*(5), 927-935. doi:10.1111/j.1469-7610.2004.t01-1-00286.x

Selig, J.P. & Little, T.D. (2012). Autoregressive and cross-lagged panel analysis for lagged panel analysis for longitudinal data. In B.P. Laursen, T.D. Little, & N.A. Card (Eds.), *Handbook of developmental research methods* (pp.265-278). New York, NY: Guilford Press. Retrieved from: http://www.myilibrary.com.ezproxy1.lib.asu.edu?ID=343127

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society, Series B, 298* (1089),199-209.

Sherbow, A., Kettler, R.J., Elliott, S.N., Davies, M.D., & Dembitzer, L. (2015). Using the SSIS assessments with Australian students: A comparative analysis of test psychometrics to the U.S. normative sample. *School Psychology International, 36*(3), 1-9. doi:10.1177/0143034315574767.

Shields, A., & Cicchetti, D. (1997). Emotion regulation among school-age children: The development and validation of a new criterion Q-sort scale. *Developmental Psychology, 33*(6), 906-916. doi:10.1037/0012-1649.33.6.906

Silver, C. H. (2014). Sources of data about children’s executive functioning: Review and commentary. *Child Neuropsychology, 20*(1), 1-13. doi:10.1080/09297049.2012.727793

Smolker, H. R., Depue, B. E., Reineberg, A. E., Orr, J. M., & Banich, M. T. (2015). Individual differences in regional prefrontal gray matter morphometry and fractional anisotropy are associated with different constructs of executive function. *Brain Structure and Function, 220*(3), 1291-1306. doi:10.1007/s00429-014-0723-y

Somsen, R. J. M., Van der Molen, Maurits W, Richard Jennings, J., & van Beek, B. (2000). Wisconsin card sorting in adolescents: Analysis of performance, response times and heart rate. *Acta Psychologica, 104*(2), 227-257. doi:10.1016/S0001-6918(00)00030-5

Sørensen, L., Plessen, K. J., & Lundervold, A. J. (2012). The impact of inattention and emotional problems on cognitive control in primary school children. *Journal of Attention Disorders, 16*(7), 589-599. doi:10.1177/1087054711417394

Sparrow, S., Cicchetti, D., & Balla, D. (2005). *Vineland Adaptive Behavior Scales: Second Edition.* Minneapolis, MN: Pearson Assessments.
Statistics Canada. (1995). *National Longitudinal Survey of Children: Survey instruments for 1994-95*. Data collection cycle1 (No. Cat. No. 95-01). Ottawa: Statistics Canada.

Staub, E. (1978). *Positive social behavior and morality: Volume 1; social and personal influences*. New York, NY: Academic Press.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology, 59*(4), 745-759. doi:10.1080/17470210500162854

St Clair-Thompson, H., Stevens, R., Hunt, A., & Bolder, E. (2010). Improving children's working memory and classroom performance. *Educational Psychology, 30*(2), 203-219. doi:10.1080/01443410903509259

Stichter, J. P., Christ, S. E., Herzog, M. J., O'Donnell, R. M., & O'Connor, K. V. (2016). Exploring the role of executive functioning measures for social competence research. *Assessment for Effective Intervention, 41*(4), 243-254. doi:10.1177/1534508416644179

Stipek, D., Newton, S., & Chudgar, A. (2010). Learning-related behaviors and literacy achievement in elementary school-aged children. *Early Childhood Research Quarterly, 25*(3), 385-395.

Stone, L. L., Otten, R., Engels, R. C. M. E., Vermulst, A. A., & Janssens, J. M. A. M. (2010). Psychometric properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4-to 12-year-olds: A review. *Clinical Child and Family Psychology Review, 13*(3), 254-274. doi:10.1007/s10567-010-0071-2

Strayer, J., & Roberts, W. (1997). Facial and verbal measures of Children's emotions and empathy. *International Journal of Behavioral Development, 20*(4), 627-649. doi:10.1080/016502597385090

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643-662. doi:10.1037/h0054651

Swanson, H. L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 96*(3), 471-491. doi:10.1037/0022-0663.96.3.471

Swanson, H. L., Cooney, J. B., & McNamara, J. K. (2004). Memory and learning disabilities. In B. Y. Wong (Ed.), *Understanding Learning Disabilities* (3rd ed., pp. 41–92). San Diego, CA: Academic Press.

Swanson, H.L., Zheng, X, & Jerman, O. (2009). Working Memory, Short-Term Memory, and Reading Disabilities: A Selective Meta-Analysis of the Literature. *Journal of Learning Disabilities*, *42*(3), 260–287.https://doi.org/10.1177/0022219409331958

Tabachnick, B., & Fidell, L. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson/Allyn & Bacon.

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development, 88*(4), 1156-1171. doi:10.1111/cdev.12864

Thorell, L. B. & Catale, C. (2014). The assessment of executive functioning using the Childhood Executive Functioning Inventory (CHEXI). In S. Goldstein & J.A. Naglieri (Eds.) *Handbook of executive functioning* (pp. 359-366). New York, NY: Springer. doi:10.1007/978-1-4614-8106-5

Thorell, L. B., Eninger, L., Brocki, K. C., & Bohlin, G. (2010). Childhood executive function inventory (CHEXI): A promising measure for identifying young children with ADHD? *Journal of Clinical and Experimental Neuropsychology, 32*(1), 38-43. doi:10.1080/13803390902806527

Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, *12*(1), 106-113. doi:10.1111/j.1467-7687.2008.00745.x

Thorell, L. B., & Nyberg, L. (2008). The childhood executive functioning inventory (CHEXI): A new rating instrument for parents and teachers. *Developmental Neuropsychology, 33*(4), 536-552. doi:10.1080/87565640802101516

Thorell, L. B., Veleiro, A., Siu, A. F. Y., & Mohammadi, H. (2013). Examining the relation between ratings of executive functioning and academic achievement: Findings from a cross-cultural study. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence, 19*(6), 630-638. doi.org/10.1080/09297049.2012.727792

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scales*. Chicago, IL: Riverside Publishing.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct?: Performance-based and rating measures of EF. *Journal of Child Psychology and Psychiatry, 54*(2), 131-143. doi:10.1111/jcpp.12001

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). TOWRE: Test of word reading efficiency. Austin, TX: Pro-Ed.

Torrente, C., Alimchandani, A., & Aber, J.L. (2015). International perspectives on SEL. In J.A. Durlak (Ed.), *Handbook of social and emotional learning: Research and practice* (pp. 566-587). New York, NY: The Guilford Press. Retrieved from http://search.ebscohost.com.ezproxy.library.ubc.ca/login. aspx?direct=true&db=nlebk&AN=969923&site=ehost-live&scope=site

Tourangeau, K., Nord, C., Le, T., Sorongon, A. G., & Najarian, M. (2009). Early Childhood Longitudinal Study, Kindergarten class of 1998–199 (ECLS-K), combined user's manual for the ECLS-K eighth-grade and K-8 full sample data files and electronic codebooks (NCES 2009-004). Retrieved from National Center for Education Statistics, Institute of Education Sciences (website: http://nces.ed.gov/ecls/data/ECLSK_K8_Manual_part1.pdf).

Traverso, L., Viterbori, P., & Usai, M. C. (2015). Improving executive function in childhood: Evaluation of a training intervention for 5-year-old children. *Frontiers in Psychology, 6*(525), 1-14. doi:10.3389/fpsyg.2015.00525

Tremblay, R. E., Loeber, R., Gagnon, C., Charlebois, P., Larivée, S., & LeBlanc, M. (1991). Disruptive boys with stable and unstable high fighting behavior patterns during junior elementary school. *Journal of Abnormal Child Psychology, 19*(3), 285-300. doi:10.1007/BF00911232

Tudge, J. R. H., Mokrova, I., Hatfield, B. E., & Karnik, R. B. (2009). Uses and misuses of Bronfenbrenner's bioecological theory of human development. *Journal of Family Theory & Review*, *1*(4), 198-210. doi:10.1111/j.1756-2589.2009.00026.x

U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD) (June, 2018). Table 204.10. "Public Elementary/Secondary School Universe Survey," 2000-01, 2010-11, 2014-15, and 2015-16. https://nces.ed.gov/programs/digest/d17/tables/dt17_204.10.asp

Van Boxtel, G. J. M., Van der Molen, M. W., Jennings, J. R., & Brunia, C.H. M. (2001). A psychophysiological analysis of inhibitory motor control in the stop-signal paradigm. *Biological Psychology*, *58*(3), 229-262. doi:10.1016/S0301-0511(01)00117-X

van der Sluis, S., van der Leij, A., & de Jong, P. F. (2005). Working memory in dutch children with reading-and arithmetic-related LD. *Journal of Learning Disabilities, 38*(3), 207-221. doi:10.1177/00222194050380030301

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence, 35*(5), 427-449. doi:10.1016/j.intell.2006.09.001

Van der Ven, Sanne H G, Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *The British Journal of Educational Psychology, 82*(1), 100.

Van der Ven, S. H. G, Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *The British Journal of Educational Psychology, 82*(1), 100-119. doi:10.1111/j.2044-8279.2011.02035.x

Vera-Estay, E., Seni, A. G., Champagne, C., & Beauchamp, M. H. (2016). All for one: Contributions of age, socioeconomic factors, executive functioning, and social cognition to moral reasoning in childhood. *Frontiers in Psychology, 7*, 227-240. doi:10.3389/fpsyg.2016.00227

Vuchinich, S., Flay, B . R., Aber, L., & Bickman, L. (2012). Person mobility in the design and analysis of cluster-randomized cohort prevention trials. *Prevention Science, 13*(3), 300-313. doi: 10.1007/s11121-011-0265-y

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin, 140*(4), 1174-1204. doi:10.1037/a0036620

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner & E. Souberman., Eds.) (A. R.

Luria, M. Lopez-Morillas & M. Cole [with J. V. Wertsch], Trans.) Cambridge, Mass.: Harvard University Press. (Original manuscripts [ca. 1930- 1934])

Waber, D. P., Gerber, E. B., Turcios, V. Y., Wagner, E. R., & Forbes, P. W. (2006). Executive functions and performance on high-stakes testing in children from urban schools. *Developmental Neuropsychology, 29*(3), 459-477. doi:10.1207/s15326942dn2903_5

Weber, R. C., Johnson, A., Riccio, C. A., & Liew, J. (2016). Balanced bilingualism and executive functioning in children. *Bilingualism: Language and Cognition, 19*(2), 1-7. doi:10.1017/S1366728915000553

Wechsler, D. (1981). *WAIS-R manual*. New York, NY: Psychological Corporation.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children* (3rd ed., Psykologiförlaget AB, Stockholm, trans.). New York, NY: Psychological Corporation.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Harcourt Assessment.

Wechsler, D. (2001). Wechsler Individual Achievement Test (Vol. 2). San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2003) *Wechsler Intelligence Scale for Children - Fourth Edition (WISC-IV) administration and scoring manual*. San Antonio, TX: The Psychological Corporation.

Weintraub, S., Bauer, P. J., Zelazo, P. D., Wallner-Allen, K., Dikmen, S. S., Heaton, R. K., . . . Gershon, R. C. (2013). i. NIH toolbox cognition battery (CB): Introduction and pediatric data: *Monographs of the Society for Research in Child Development, 78*(4), 1-15. doi:10.1111/mono.12031

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology, 108*(3), 436-452. doi:10.1016/j.jecp.2010.08.008

Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development, 84*(6), 2112-2130. doi:10.1111/cdev.12099

Weir, K., & Duveen, G. (1981). Further development and validation of the prosocial behaviour questionnaire for use by teachers. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 22*(4), 357-374. doi:10.1111/j.1469-7610.1981.tb00561.x

Weissberg, R. P., Durlak, J. A., Domitrovich, C. E., & Gullotta, T. P. (2015). Social and emotional learning: Past, present, and future. In J.A. Durlak (Ed.), *Handbook of social and emotional learning: Research and practice* (pp. 3-19). New York, NY: The Guilford Press. Retrieved from http://search.ebscohost.com.ezproxy.library.ubc.ca/login.aspx?direct=true&db=nlebk&AN=969923&site=ehost-live&scope=site

Welsh, M. C. (2001). Prefrontal cortex and the development of executive function in childhood. In A.F. Kalverboer & A.A. Gramsbergen (Eds.) *Handbook of brain and behaviour in human development* (pp. 767–789). Norwell, MA: Kluwer Academic Publishers.

Welsh, M. C., Friedman, S. L., & Spieker, S. J. (2006). Executive functions in developing children: Current conceptualizations and questions for the future. In K. McCartney & D. A. Phillips (Eds.) *Blackwell Handbook of Early Childhood Development* (pp. 167-187). Malden, MA: Blackwell Pub. doi:10.1002/9780470757703.ch9

Welsh, M. C., & Huizinga, M. (2001). The development and preliminary validation of the Tower of Hanoi-revised. *Assessment, 8*(2), 167-176. doi:10.1177/107319110100800205

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology, 102*(1), 43–53. doi:10.1037/a0016738

Welsh, M., Parke, R.D., Widaman, K., & O'Neil, R. (2001). Linkages between children's social and academic competence: A longitudinal analysis. *Journal of School Psychology, 39*(6), 463-482. doi:10.1016/S0022-4405(01)00084-X

Welsh, M. C., & Pennington, B. F. (1988). Assessing frontal lobe functioning in children: Views from developmental psychology. *Developmental neuropsychology*, *4*(3), 199-230.

Welsh, M., & Peterson, E. (2014). Issues in the conceptualization and assessment of hot executive functions in childhood. *Journal of the International Neuropsychological Society, 20*(2), 152-156. doi:10.1017/S1355617713001379

Wentzel, K. R., & Caldwell, K. (1997). Friendships, peer acceptance, and group membership: Relations to academic achievement in middle school. *Child Development, 68*(6), 1198-1209. doi:10.1111/j.1467-8624.1997.tb01994.x

Wigelsworth, M., Lendrum, A., Oldfield, J., Scott, A., ten Bokkel, I., Tate, K., & Emery, C. (2016). The impact of trial stage, developer involvement and international transferability on universal social and emotional learning programme outcomes: A meta-analysis. *Cambridge Journal of Education, 46*(3), 347-376. doi:10.1080/0305764X.2016.1195791

Wilson, B. A., Alderman, N., Burgess, P., Emslie, H., & Evans, J. J. (1996). *Behavioural assessment of the dysexecutive syndrome (BADS)*. Bury St. Edmunds, England: Thames Valley Test Company.

Wispé, L. (1972). Positive Forms of Social Behavior: An Overview. *Journal of Social Issues,28*(3), 1-19. Retrieved from https://search.lib.asu.edu/primo-explore/fulldisplay?docid=TN_wj10.1111/j.1540-4560.1972.tb00029.x&context=PC&vid=01ASU&search_scope=Everything&tab=default_tab&lang= en_US

Wolfe, K. R., Vannatta, K., Nelin, M. A., & Yeates, K. O. (2015). Executive functions, social information processing, and social adjustment in young children born with very low birth weight. *Child Neuropsychology, 21*(1), 41-54. doi:10.1080/09297049.2013.866217

Wolfe, K. R., Walsh, K. S., Reynolds, N. C., Mitchell, F., Reddy, A. T., Paltin, I., & Madan-Swain, A. (2013). Executive functions and social skills in survivors of pediatric brain tumor. *Child Neuropsychology, 19*(4), 370-384. doi:10.1080/09297049.2012.669470

Woodcock, R. W., & Johnson, M. B. (1989-1990). Woodcock–Johnson Psycho–Educational Battery–Revised. Allen,TX: DLM Teaching Resources.

Woodcock, R. W., & Mather, N. (2000). *Woodcock–Johnson Psycho-educational Battery-III*. Itasca, IL: Riverside.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson-III tests of cognitive abilities*. Itasca, IL: Riverside.

Wright, I., Waterman, M., Prescott, H., & Murdoch-Eaton, D. (2003). A new Stroop-like measure of inhibitory function development: Typical developmental trends. *Journal of Child Psychology and Psychiatry, 44*(4), 561-575. doi:10.1111/1469-7610.00145

Yeates, K. O. (2015). Social problem-solving and social adjustment in paediatric traumatic brain injury. *Brain Injury, 29*(13-14), 1682-1690. doi:10.3109/02699052.2015.1075140

Yeates, K. O. (2013). Social outcomes in pediatric traumatic brain injury: Perspectives from social neuroscience and developmental psychology. *Journal of the International Neuropsychological Society, 19*(5), 493-496. doi:10.1017/S1355617713000398

Yeates, K. O., Bigler, E. D., Dennis, M., Gerhardt, C. A., Rubin, K. H., Stancin, T., . . . Vannatta, K. (2007). Social outcomes in childhood brain disorder: A heuristic integration of social neuroscience and developmental psychology. *Psychological Bulletin, 133*(3), 535-556. doi:10.1037/0033-2909.133.3.535

Yeniad, N., Malda, M., Mesman, J., van IJzendoorn, M. H., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. *Learning and Individual Differences*, *23*, 1-9. doi:10.1016/j.lindif.2012.10.004

Yeniceri, N., & Altan-Atalay, A. (2011). Age-related changes in the Wisconsin card sorting test performances of 8- to 11-year-old Turkish children. *The Clinical Neuropsychologist*, *25*(7), 1179-1192. doi:10.1080/13854046.2011.613855

Yu, C. Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes* (Doctoral dissertation). Available from ProQuest Dissertations & Theses Global (Order No. 276287121). Retrieved from http://login.ezproxy1.lib.asu.edu/login?url=https://search-proquest-com.ezproxy1.lib.asu.edu/docview/276287121?accountid=4485

Yu, C.-Y., & Muthen, B. (2002). Evaluation of model fit indices for latent variable models with categorical and continuous outcomes. In annual meeting of the American Educational Research Association, New Orleans, LA.

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, *1*(1), 297-301. doi:10.1038/nprot.2006.46

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

Zelazo, P.D., Carlson, S.M., & Kesek, A. (2014). The development of executive function in childhood. In C.A. Nelson & M.L. Collins (Eds.), *Handbook of developmental cognitive neuroscience* (2nd ed.). Cambridge, MA: The MIT Press.

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development, 11*(1), 37-63. doi:10.1016/S0885-2014(96)90027-1

Zelazo, P.D., Jacques, S., Burack, J. A., & Frye, D. (2002). The relation between theory of mind and rule use: Evidence from persons with autism-spectrum disorders. *Infant and Child Development, 11*(2), 171-195. doi:10.1002/icd.304

Zelazo, D. P. & Muller, U. (2010). Executive function in typical and atypical development. In U. Goswami (Ed.), Wiley-*Blackwell handbook of childhood cognitive development* (2nd ed., pp. 574-603). Oxford, UK: Wiley-Blackwell. doi:10.1002/9780470996652.ch20

Zhou, Q., Chen, S. H., & Main, A. (2012). Commonalities and differences in the research on children's effortful control and executive function: A call for an integrated model of self-regulation. *Child Development Perspectives, 6*(2), 112. doi:10.1111/j.1750-8606.2011.00176.x

Zimmer-Gembeck, M. J., Geiger, T. C., & Crick, N. R. (2005). Relational and physical aggression, prosocial behavior, and peer relations: Gender moderation and bidirectional associations. *The Journal of Early Adolescence, 25*(4), 421-452. doi:10.1177/0272431605279841

Zimmermann, P., Gondan, M., & Fimm, B. (2004). Testbatterie zur Aufmerksamkeitsprüfung für Kinder (KITAP) [Battery of Tests for Assessing Attention in Children (KITAP)]. Herzogenrath, Germany: Psytest.

Zins, J.E., Weissberg, R.P., Wang M.C. & Walberg, H.J. (2004a), Acknowledgements. In J.E. Zins, R.P. Weissberg, M.C. Wang & H.J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say?* (pp. ix - x). New York, NY: Teachers College Press.

Zins, J.E., Bloodworth, M.R., Weissberg, R.P., & Walberg, H.J. (2004b). The scientific base linking social and emotional learning to school success. In J.E. Zins, R.P. Weissberg, M.C. Wang & H.J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say?* (pp. 3-22). New York, NY: Teachers College Press.

APPENDIX A

TABLES

Table A1

Executive Functioning Tests Cited and Reviewed

EF Tests in Appendix C	EF Tests Cited In Literature Review
Working Memory (i.e., updating)	
"*Store and manipulate information over brief periods of time*" *(Alloway et al., 2006, p.1698)*	
Digit Span Backwards/ Backward Word Span Wechsler Intelligence Scale for Children – 4th edition (Wechsler, 2003) *Description:* repeat sequence of numbers or letters in backwards order.	Backwards Word Span (Davis & Pratt, 1996) Counting Span (Case et al., 1982) Digit Span Backwards (Alloway, 2007) Digit Span WISC – III Variation (Wechsler, 1991) Digit Span Test (Wechsler, 1991) Digit Span Tests (Eilertsen & Johnsen, 2003) Dual Task Performance (Baddeley et al., 1997) Keep Track (Van der Sluis et al., 2007) Letter-Number Sequencing (WISC-IV; Wechsler, 2003) Listening Span Test (van der Sluis et al., 2005) Odd One Out (Alloway, 2007) Stanford Binet Intelligence (Thorndike et al., 1986) WAIS-R-NI (Wechsler, 1981) Working Memory Test Battery for Children (Pickering & Gathercole, 2001)
Inhibition (i.e., response inhibition, inhibitory control, self-restraint, self-stopping)	
(a) Suppress or disrupt a dominant response; *(b) stop behavior towards a goal or an ongoing response that is inefficient; and* *(c) resist distractions/interference to persist towards the goal*	
Stroop Color-Word Interference Test (Stroop, 1935) *Description*: Reading colored words in a non-corresponding Colored ink. Scores based on timed reading of color-word combinations	Animal Stroop (Wright et al., 2003) Day/Night Stroop Task Variations (Gerstadt et al., 1994; Berlin & Bohlin, 2002) Eriksen Flanker Task (Eriksen & Eriksen, 1974) Flanker Task Variation (Ridderinkhof & van der Molen. 1995) Go/No Go Variaton (Berlin & Bohlin, 2002) Stop-Signal (van Boxtel et al., 2001) Wisconsin Card Sorting Task (Heaton et al., 1993) Wisconsin Card Sorting Task Computerized Version (Somsen et al., 2000)

Table A1 (continued)

EF Tests Reviewed in Appendix C	EF Tests Cited In-Text
Shifting (i.e., (cognitive) flexibility, attentional set-shifting, attention switching, task switching)	
"Freely moving from one situation, activity, or aspect of a problem to another as the situation demands" (Gioia et al., 2000, p.321).	
Wisconsin Card-Sorting Test (Berg, 1948) *Description: Time and number of errors made when figuring out the rules to sort cards by color, shape, then number.*	Advanced Dimensional Change Card Sort (Zelazo et al., 1996) Children Coloured Trail Test (Llorente et al., 2003) Cognitive Flexibility Task (Zimmermann et al., 2004) Color Trail Test (D'Elia et al., 1994) Dimensional Change Card Sort (Zelazo et al., 2006) Delis-Kaplan Executive Function System (Delis et al., 2001) Wisconsin Card Sorting Task Computer Version (Somsen et al., 2000) Dots Tasks (Diamond et al., 2007) Go/No Go Accuracy and Time Subtests (Berlin & Bohlin, 2002)
Planning/Problem Solving	
"Ability to plan actions in advance and to approach a task in an organized, strategic and efficient manner" (Anderson, 2002).	
Tower of Hanoi (Piaget, 1974/76) Tower of London (Shallice, 1982) *Description: moving objects onto pegs according to a set of rules and within a certain number of moves.*	n/a
Emotional Regulation	
"Regulating emotional responses such that the emotion conveyed is appropriate to the situation or stressor (Gioia et al., 2000) for the purpose of facilitating adaptive functioning (Dodge & Garber, 1991).	
Children's Gambling Task (Kerr & Zelazo, 2004) *Description: cards chosen from two decks, one of which is riskier. Happy/sad faces determine wins/loses.*	n/a

Table A2

Overlap Among Rating Scale Items

EXECUTIVE FUNCTIONING	SOCIAL EMOTIONAL LEARNING SKILLS
WORKING MEMORY	
Remember	
When shown something complicated to do he/she cannot keep it in mind so as to do it correctly (BDEFS – CA, 14)	Remember important information (DESSA – SSE, 1)
Difficulty remembering lengthy instructions (CHEXI, 1)	
Steps of a Task	
Has trouble doing things...in proper order (BDEFS–CA, 24)	Perform the steps of a task in order (DESSA – SSE, 23)
When asked to do several things he/she only remembers the first or last (CHEXI, 6)	
Has difficulty carrying out activities that require several steps (CHEXI, 14)	
Has difficulty with tasks/activities with several steps (CHEXI, 20)	
Learn from Experience	
Unable to “think on his/her feet”, problem solve, or respond effecitvely to unexpected events (BDEFS–CA, 25)	Learns from experience (DESSA – SSE, 30)
Difficulty thinking ahead/learning from experience (CHEXI, 21)	

Table A2 (continued)

EXECUTIVE FUNCTIONING	SOCIAL EMOTIONAL LEARNING SKILLS
INHIBITION	
Disrupts Response/Resist Interference	
Has difficulty waiting for things; has to have things or do things he/she wants right away (BDEFS – CA, 28)	Wait for turn (DESSA – SSE, 21)
	Pass up something he/she wanted, or do something he/she didn't like to get something better in future (DESSA – SSE, 25)
Stop Behavior	
Unable to inhibit his/her reactions to events or to what others say or do to him/her; reacts on impulse (BDEFS – CA, 30)	Thinks things out before acting (SDQ, 21)
Likely to do things without considering the consequences (BDEFS – CA, 34)	Acts without thinking (SSIS – RS/ SSIS – SEL RF, 47)
Acts without thinking things over (BDEFS – CA, 35)	
Adjust to Change	
Is disturbed by changes in teacher of class (BRIEF – 2, 60)	Adjusts well to changes in plans (DESSA – SSE,31)
	Adjusts well when going from one setting to another (DESSA – SSE, 36)

Table A2 (continued)

EXECUTIVE FUNCTIONING	SOCIAL EMOTIONAL LEARNING SKILLS
SHIFTING (continued)	
Getting Stuck	
Gets stuck on one topic or activity (BRIEF – 2, 17)	Repeats the same things over and over (SSIS – RS/ SSIS – SEL RF, 60)
PLANNING/PROBLEM SOVLING	
Follow Through	
Procrastinates or puts off doing things until the last minute (BDEFS – CA, 1)	Take steps to achieve goals (DESSA – SSE, 4)
Not able to prepare in advance for things he/she knows he/she is supposed to do (BDEFS – CA, 11)	
Has difficulty planning for an activity (CHEXI, 12)	
Preparation	
Not prepared on time for schoolwork (or assigned tasks given at home (BDEFS – CA, 4)	Takes an active role in learning (DESSA – SSE, 7)
Time Management	
Has a poor sense of time (BDEFS – CA, 2)	Gets things done in a timely fashion (DESSA – SSE, 12)
Wastes or doesn’t manage his/her time well (BDEFS – CA, 3)	
Not prepared on time for schoolwork or assignemnts given at home (BDEFS – CA, 4)	
Has difficulty judging how much time it will take to do something or get somewhere (BDEFS – CA, 8)	

Table A2 (continued)

EXECUTIVE FUNCTIONING	SOCIAL EMOTIONAL LEARNING SKILLS
PLANNING/PROBLEM SOLVING (continued)	
Creativity	
Has trouble considering various ways of doing things (BDEFS – CA, 15)	Shows creativity in completing a task (DESSA – SSE, 11)
Not as creative or inventive as others of his/her age (BDEFS – CA, 20)	
Resourcefulness	
Slow at solving problems he/she encounters in his/her daily life (BDEFS – CA, 26)	Seeks out additional knowledge or information (DESSA – SSE, 6)
	Use available resources (people or objects) to solve a problem (DESSA – SSE, 33)
EMOTIONAL REGULATION	
Stays Calm	
Reacts more strongly to situations than other children (BRIEF – 2, 27)	Stays calm when faced with a challenge (DESSA – SSE, 28)
Tantrums	
Finds it difficult to walk away from emotionally upsetting encounters with others or leave situations in which he/she becomes very emotional (BDEFS – CA, 66)	Often has temper tantrums or hot temper (SDQ, 5)
Remains emotional or upset longer than other children (BDEFS – CA, 65)	Has temper tantrums (SSIS – RS, 57)

Table A2 (continued)

EXECUTIVE FUNCTIONING	SOCIAL EMOTIONAL LEARNING SKILLS
EMOTIONAL REGULATION (continued)	
Other Various Strong Emotions	
Quick to get angry or become upset (BDEFS – CA, 57) Overreacts emotionally (BDEFS – CA, 58)	Many worries, often seems worried (SDQ, 8) Many fears, easily scared (SDQ, 24) Often unhappy, down-hearted, or tearful (SDQ, 13)

Note. BDEFS – CA = Barkley Deficits in Executive Functioning Scale – Children and Adolescents (Barkley, 2012b); BRIEF – 2 = Behavior Rating Inventory of Executive Function Reproduction by special permission of the Publisher, Psychological Assessment Resources, Inc. (PAR), 16204 North Florida Avenue, Lutz, Florida, 33459 from the Behavior Rating Inventory of Executive Function, Second Edition by Gerald A. Gioia, PhD, Peter K. Isquith, PhD, Stephen C. Guy, PhD, and Lauren Kenworthy, PhD, Copyright 1996, 1998, 2000, 2015 by PAR. Further reproduction is prohibited without permission from PAR; CHEXI = Childhood Executive Functioning Inventory (Thorell & Nyberg, 2008); DESSA – SSE = Devereux Student Strength Assessment – Second Step® Edition (Devereux Center for Resilient Children, 2012); SDQ = Strengths and Difficulties Questionnaire (Goodman, 1997); SSIS – RS = Social Skills Improvement System Rating Scales (Gresham & Elliott, 2008); SSIS SEL – RF = Social Skills Improvement System SEL Edition Rating Form (Gresham & Elliott, 2017).

Table A3

Descriptive Statistics of the Student Sample by Cohort

Variable	Cohort 1 *n* (%)		Cohort 2 *n* (%)	
Grade				
Kindergarten	1435 (100%)		-	
Grade 1	-		1594 (100%)	
Missing	0		0	
Gender				
Male	551 (38.4)		704 (44.2)	
Female	546 (38.0)		724 (45.4)	
Missing	338 (23.6)		166 (10.4)	
Ethnicity				
Asian	179 (12.5)		171 (10.7)	
American Indian/ Alaska Native	39 (2.7)		93 (5.8)	
Black/ African American	102 (7.1)		106 (6.0)	
Caucasian/White	584 (40.7)		563 (35.3)	
Hispanic	324 (22.6)		476 (29.9)	
Native Hawaiian or Other Asian /Pacific Islander	25 (1.7)		45 (2.8)	
More than one ethnicity	96 (6.7)		131 (8.2)	
Missing	86 (6.0)		79 (5.0)	
English Language Learner (ELL) status				
Not an ELL	756 (52.7)		854 (53.6)	
ELL Student	255 (17.8)		415 (26.0)	
Missing	424 (29.5)		325 (20.4)	
Lunch program distribution				
24% or less	140 (9.8)		118 (7.4)	
25% - 49%	403 (28.1)		413 (25.9)	
50% - 74%	492 (34.3)		506 (31.7)	
75% or more	400 (27.9)		557 (34.9)	
Missing	0		0	
Location				
Arizona	332 (23.1)		486 (30.5)	
Washington	774 (53.9)		982 (61.6)	
Missing	329 (22.9)		126 (7.9)	
	Year 1	Year 2	Year 1	Year 2
Age				
M (SD)	5.6 (.354)	6.6 (.345)	6.6 (.345)	7.6 (.378)
Range (years)	4.8 – 8.1	5.8 – 9.1	5.2 – 8.1	6.1 – 9.0
Missing (%)	0	0	0	0

Table A4

Descriptive Statistics of the Teacher Sample (Fall Year 1)

Variable	*n*	%
Gender		
Male	9	6.0
Female	142	94.0
Missing	0	0
Hispanic or Latino/a		
No	142	94
Yes	9	6
Missing	0	0
Ethnicity		
Asian	6	4.0
American Indian/ Alaska Native	1	0.7
Black/ African American	0	0
Caucasian/White	128	84.8
Native Hawaiian or Other Asian/Pacific Islander	0	0
More than one ethnicity	10	6.6
Other	6	4.0
Missing	0	0
Grade(s) taught		
Kindergarten	61	40.4
Kindergarten / Grade 1 Split	4	2.7
Grade 1	75	49.7
Grade 1 / Grade 2 Split	4	2.7
Missing	7	4.6
Highest degree received		
Bachelor's degree	48	31.8
Master's degree	87	57.6
Professional degree	6	4.0
Doctorate degree	1	0.7
Missing	9	6.0
Location		
Arizona	48	31.8
Washington	103	68.2
Missing	0	0

	n	*M (SD)*
Age	149	42.9 (11.9)
Missing	2	
Number of years teaching	151	14.4 (9.4)
Missing	0	

Table A5

Current Study Measures by Factor

Factor 1 Executive Functioning 1 (EF 1)
Working Memory Component Items
Remember important information (DESSA – SSE 1)
Performs steps of task in order (DESSA – SSE 23)
Learns from experience (DESSA – SSE 30)
Plan/Problem Solving Component Items
Take steps to achieve goals (DESSA – SSE 4)
Takes an active role in learning (DESSA – SSE 7)
Gets things done in timely fashion (DESSA – SSE 12)
Shows creativity in completing a task (DESSA–SSE 11)
Seeks out additional knowledge/information (DESSA – SSE 6)
Use available resources to solve a problem (DESSA – SSE 33)
Factor 2 Executive Functioning 2 (EF 2)
Shifting Component Items
Adjusts well to changes in plans (DESSA – SSE 31)
Adjusts well when going from one setting to another (DESSA – SSE 36)
Emotional Regulation Component Items
Often loses temper (SDQ 5; Reversed)
Often unhappy, depressed, or fearful (SDQ 13; Reversed)

Table A5 (continued)

Factor 3 Prosocial Skills (PS)
Considerate of other people's feelings (SDQ 1)
Shares readily with other children. For example, toys, treats, pencils (SDQ 4)
Helpful is someone is hurt, upset, or feeling ill (SDQ 9)
Kind to younger children (SDQ 17)
Often offers to help others (parents, teachers, other children) (SDQ 20)

Note. The items above were derived from two SEL Skills scales, abbreviated as DESSA – SSE for the Devereux Student Strengths Assessment – Second Step® Edition (Devereux Center for Resilient Children, 2012), and the SDQ for the Strengths and Difficulties Questionnaire (Goodman, 1997). Item numbers follow the rating scale abbreviations.

Table A6

Descriptive Statistics of Non-Academic Variables by Cohort

Measure	*N*	% Missing	*M*	Range Observed	Range Potential	*SD*	Skew	Kurtosis	Alpha
EF 1									
Cohort 1									
T1	1,011	29.5	2.66	0.22–4	0–4	0.78	-0.40	-0.07	.95
T2	982	31.6	2.96	0.22–4	0–4	0.74	-0.47	-0.15	.95
T3	1,174	18.2	2.79	0–4	0–4	0.76	-0.38	-0.13	.95
T4	1,019	22.0	2.87	0.11– 4	0–4	0.78	-0.42	-0.30	.95
Cohort 2									
T1	1,326	16.8	2.70	0–4	0–4	0.78	-0.38	0.04	.95
T2	1,311	17.8	2.86	0–4	0–4	0.78	-0.43	-0.26	.95
T3	1,154	27.6	2.82	0.67–4	0–4	0.74	-0.31	-0.31	.95
T4	1,113	30.2	2.88	0.56–4	0–4	0.75	-0.29	-0.49	.95
EF 2									
Cohort 1									
T1	1,011	29.5	1.80	0.25–3	0–3	0.48	-1.32	2.40	.73
T2	983	31.5	1.80	0.25–3	0–3	0.45	-0.98	1.20	.71
T3	1,175	18.1	2.45	0.50–3	0–3	0.47	-1.10	1.56	.75
T4	1,119	22.0	2.46	0–3	0–3	0.50	-1.26	2.15	.77
Cohort 2									
T1	1,331	16.5	2.44	0.50–3	0–3	0.48	-1.08	1.30	.74
T2	1,311	17.8	2.48	0–3	0–3	0.50	-1.42	2.93	.78
T3	1,167	26.8	2.41	0.50–3	0–3	0.49	-1.01	1.09	.74
T4	1,114	30.1	2.44	0.25–3	0–3	0.52	-1.07	0.95	.78
Prosocial Skills									
Cohort 1									
T1	1,010	29.6	1.39	0–2	0–2	0.49	-0.42	-0.66	.83
T2	984	31.4	1.56	0–2	0–2	0.46	-0.92	0.22	.83
T3	1,168	18.6	1.55	0–2	0–2	0.45	-0.76	-0.24	.84
T4	1,118	22.1	1.56	0–2	0–2	0.46	-0.99	0.38	.84
Cohort 2									
T1	1313	17.6	1.46	0–2	0–2	0.48	-0.53	-0.59	.85
T2	1313	17.6	1.55	0–2	0–2	0.46	-0.78	-0.28	.85
T3	1,150	27.9	1.55	0–2	0–2	0.46	-0.74	-0.38	.84
T4	1,112	30.2	1.55	0–2	0–2	0.47	-0.88	-0.06	.85

Note. EF 1= Executive Functioning (Working Memory/Plan/Problem Solving); EF 2 = Executive Functioning (Shifting/Emotional Regulation). T1 = Time 1 (fall, Y1); T2 = (spring, Y1); T3 = (fall, Y2); T4 = (spring, Y2).

Table A7

Descriptive Statistics of Academic Variables by Cohort

Measure	*N*	% Missing	*M*	Range Observed	Range Potential	*SD*	Skew	Kurtosis
Cohort 1 Reading								
T1	1,030	31.1	7.63	0–151	0–151	20.81	4.46	21.41
T2	1,030	31.1	23.69	0–155	0–155	33.70	2.14	4.21
T3	1,171	21.7	40.85	0– 221	0–242	40.00	1.34	1.18
T4	1,082	27.6	72.30	1–188	0–236	41.84	0.60	-0.52
Cohort 2 Reading								
T1	1,373	13.9	35.52	0–201	0–242	34.70	1.51	2.05
T2	1,368	14.2	68.08	0–184	0–236	38.17	0.68	-0.23
T3	1,252	21.5	79.76	1–206	0–314	39.78	0.25	-0.44
T4	1,152	27.7	96.58	6–225	0–303	36.07	0.36	-0.02
Cohort 1 Math								
T1	1,035	30.8	5.07	0–100	0–100	10.23	4.67	30.83
T2	1,035	30.8	26.39	0–100	0–100	21.01	1.16	1.42
T3	1,167	21.9	53.63	0–100	0–100	23.28	-0.24	-0.51
T4	1,084	27.5	79.30	0–100	0–100	18.21	1.13	1.03
Cohort 2 Math								
T1	1,380	13.4	45.18	0–100	0–100	22.86	-0.01	-0.57
T2	1,367	14.2	78.70	3.6–100	0–100	17.57	-1.01	1.09
T3	1,256	21.2	61.81	0–100	0–100	18.55	-0.30	-0.32
T4	1,132	29.0	81.81	17.9–100	0–100	15.82	-1.16	1.19

Note. Reading = Reading Curriculum-Based Measure (CBM); words correct per minute.

Math = CBM, percentage correct. T1 = Time 1 (fall, Y1); T2 = (spring, Y1); T3 = (fall, Y2);

T4 = (spring, Y2).

Table A8

Correlations Among Executive Functioning (EF 1 and EF 2) and Prosocial Skills by Cohort

	EF 1 T1	EF 1 T2	EF 1 T3	EF 1 T4	EF 2 T1	EF 2 T2	EF 2 T3	EF 2 T4	PS T1	PS T2	PS T3	PS T4
EF 1 T1	–	.74**	.44**	.46**	.64**	.52**	.33**	.35**	.64**	.47**	.25**	.28**
EF 1 T2	.66**	–	.49**	.50**	.53**	.70**	.34**	.37**	.51**	.58**	.22**	.26**
EF 1 T3	.41**	.46*	–	.75**	.32**	.37**	.68*	.56**	.33**	.29**	.58**	.44**
EF 1 T4	.37**	.39**	.76**	–	.29**	.39**	.54**	.71**	.37**	.34**	.50**	.58**
EF 2 T1	.60**	.47**	.26**	.25**	–	.66**	.37**	.34**	.62**	.47**	.25**	.27**
EF 2 T2	.45**	.67**	.32**	.30**	.62**	–	.41**	.41**	.48**	.58**	.27**	.32**
EF 2 T3	.27**	.30**	.70**	.53**	.27**	.33**	–	.70**	.32**	.29**	.58**	.47**
EF 2 T4	.24**	.23**	.54**	.70**	.27**	.31**	.70**	–	.37**	.35**	.52**	.61**
PS T1	.64**	.43**	.28**	.27**	.49**	.38**	.23**	.22**	–	.63**	.33**	.35**
PS T2	.46**	.58**	.30**	.26**	.40**	.53**	.29**	.26**	.63**	–	.30**	.34**
PS T3	.26**	.29**	.59**	.46**	.23**	.28**	.61**	.48**	.30**	.30**	–	.65**
PS T4	.22**	.21**	.45**	.56**	.21**	.25**	.50**	.59**	.30**	.32**	.67**	–

Note. Cohort 1 is represented on the left of the diagonal; Cohort 2 is represented on the right of the diagonal. EF 1 = Executive Functioning 1 (Working Memory/ Plan/Problem Solving) EF 2 = Executive Functioning 2 (Shifting/Emotional Regulation). PS = Prosocial Skills. T1 = Time 1; T2 = Time 2; T3 = Time 3; T4 = Time 4.

$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$; $n_{Cohort\ 1} = 713 – 1176$; $n_{Cohort\ 2} = 868 – 1332$.

Table A9

Correlations Among Academic and Non-Academic Variables by Cohort

Cohort 1	Math K		Math Gr. 1		Reading K		Reading Gr. 1	
	T1	T2	T3	T4	T1	T2	T3	T4
EF 1								
T1	.24**	.39**	.38**	.29**	.20**	.31**	.32**	.34**
T2	.24**	.35**	.45**	.36**	.20**	.31**	.34**	.38**
T3	.21**	.35**	.40**	.40**	.15**	.29**	.37**	.42**
T4	.19**	.34**	.38**	.41**	.17**	.27**	.34**	.43**
EF 2								
T1	.09**	.20**	.23**	.17**	.07*	.11**	.10**	.13**
T2	.13**	.24**	.27**	.19**	.07*	.11**	.13**	.19**
T3	.09*	.19**	.22**	.22**	.04	.11**	.17**	.21**
T4	.14**	.21**	.21**	.14**	.09*	.14**	.18**	.16**
PS								
T1	.13**	.20**	.21**	.14**	.09**	.14**	.18**	.16**
T2	.10**	.25**	.26**	.16**	.02	.08*	.10**	.14**
T3	.01	.10**	.13**	.12**	-.07	.01	.04	.08**
T4	-.02	.07**	.11**	.11**	-.04	.01	.03	.07*

Cohort 2	Math Gr.1		Math Gr. 2		Reading Gr.1		Reading Gr. 2	
	T1	T2	T3	T4	T1	T2	T3	T4
EF 1								
T1	.36**	.33**	.31*	.25**	.36**	.41**	.40**	.33**
T2	.33**	.39**	.36**	.28**	.34**	.42**	.42**	.36**
T3	.27**	.33**	.38**	.35**	.35**	.43**	.44**	.39**
T4	.25**	.32**	.33*	.34**	.31**	.38**	.40**	.39**
EF 2								
T1	.17**	.19**	.13**	.11**	.17**	.19**	.19**	.16**
T2	.17**	.27**	.21**	.16**	.17**	.25**	.23**	.18**
T3	.10**	.16**	.20**	.18**	.16**	.21**	.22**	.21**
T4	.10**	.15**	.20**	.19**	.16**	.21**	.20**	.22**
PS								
T1	.15**	.18**	.17**	.11**	.10**	.15**	.16**	.12**
T2	.12**	.14**	.12**	.10**	.06*	.12**	.13**	.09**
T3	.02	.08*	.11**	.13**	.10**	.13**	.14**	.10**
T4	.02	.06	.10**	.10**	.06	.10**	.10**	.10**

Note. EF1=Executive Functioning (working memory, plan/ problem solving); EF 2 = Shift, Emotional Regulation; PS = Prosocial Skills; K = Kindergarten; Gr.= Grade; T = Time.

* $p < .05$, **$p < .01$, ***$p < .001$; n range was 888 – 1380.

Table A10

Factor Loadings (Pattern Matrix) for 24-Item Combined Factors Model Based on Exploratory Factor Analysis with Direct Oblimin (n=1,130)

Item	Factor		
	WM/Inh/ Plan/PS	Shift/ ER	Prosocial
Hypothesized Working Memory Component of EF			
Remember important information (DESSA – SSE 1)	**1.04***	-.14*	-.18*
Performs steps of task in order (DESSA – SSE 23)	**.91***	.13*	-.10*
Learns from experience (DESSA – SSE 30)	**.71***	.14*	.12*
Hypothesized Inhibition Component of EF			
Waits his/her turn (DESSA – SSE 21)	**.43***	.38*	.20*
Passes up something he/she wanted, or does something he/she did not like to get something better in future (DESSA – SSE 25)	**.51***	.09*	.13*
Thinks things out before acting (SDQ 21)	**.35***	.24*	.35*
Hypothesized Shifting Component of EF			
Adjusts well to changes in plans (DESSA – SSE 31)	.36*	**.66***	.00
Adjusts well when going from one setting to another (DESSA – SSE 36)	.39*	**.60***	.01
Hypothesized Plan/Problem Solving Component of EF			
Take steps to achieve goals (DESSA – SSE 4)	**.95***	-.00	-.03
Takes an active role in learning (DESSA – SSE 7)	**1.10***	-.24*	-.02
Gets things done in timely fashion (DESSA – SSE 12)	**.92***	.05	-.16*
Shows creativity in completing a task (DESSA–SSE 11)	**.94***	-.17*	.02
Seeks out additional information (DESSA–SSE 6)	**1.11***	-.34*	.02

Table A10 (continued)

Item	Factor		
	WM/Inh Plan/PS	Shift/ ER	Prosocial
Hypothesized Plan/Problem Solving Component of EF			
Use available resources to solve a problem (DESSA – SSE 33)	**.76***	.11*	.08*
Hypothesized Emotional Regulation Component of EF			
Stays calm when faced with a challenge (DESSA – SSE 28)	.42*	**.45***	.07*
Often loses temper (SDQ 5; Reversed)	-.11	**.64***	.29*
Many worries, often seems worried (SDQ 8; Reversed)	.03	**.94***	-.52*
Many fears, easily scared (SDQ 24; Reversed)	.03	**.99***	-.59*
Often unhappy, depressed, or fearful (SDQ13; Reversed)	.01	**.88***	-.23*
Hypothesized Prosocial skills			
Considerate of other people's feelings (SDQ 1);	.04	.25*	**.69***
Shares readily with other children (SDQ 4)	.00	.25*	**.69***
Helpful if someone is hurt, upset, or feeling ill (SDQ 9);	.04	.04	**.78***
Kind to young children (SDQ 17)	.01	.18*	**.67***
Often offers to help others (SDQ 20)	.19*	-.01	**.70**

Note. DESSA – SSE = Devereux Student Strengths Assessment – Second Step®
Edition (Devereux Center for Resilient Children, 2012); SDQ = Strengths and
Difficulties Questionnaire (Goodman, 1997).

* $p < .05$.

Table A11

Fit Indices for 24-Item Combined Factors Model Based on Exploratory Factor Analyses with Direct Oblimin (n = 1,130)

Model	χ^2	*df*	RMSEA	[90% CI]	WRMR	CFI
1-Factor	6707.19**	252	.151	[.147, .154]	.107	.922
2-Factor	4199.96**	229	.124	[.121, .127]	.089	.952
3-Factor	2372.84**	207	.096	[.093, .100]	.057	.974
4-Factor	1311.88**	186	.073	[.069, .077]	.031	.986

Note. RMSEA = Root Mean Square Error of Approximation; 90% CI: 90% Confidence Interval for RMSEA; WRMR = Weighted Root Mean Square Residual; CFI = Comparative Fit Index.

**$p < 0.01$

Table A12

Fit Indices by Model Based on Longitudinal Confirmatory Factor Analysis (n = 2,931)

	χ^2			χ^2 DIFF Test			Fit Indices				
Model	χ^2	*df*	*p*	χ^2	*df*	*p*	RMSEA	[90% CI]	WRMR	CFI	ΔCFI Pass/ Fail
Executive Functioning Model EF 1 (Working Memory/Plan/Problem Solving)											
1.	1674.01	132	<.001	–	–	–	.063	.060, .066	0.029	.989	–
2.	1705.77	140	<.001	28.55	8	<.001	.063	.059, .064	0.029	.989	0 Pass
3.	1784.23	166	<.001	115.10	26	<.001	.058	.055, .060	0.029	.988	.001 Pass
4.	1734.57	175	<.001	64.01	9	<.001	.055	.053, .058	0.030	.989	.001 Pass
Executive Functioning Model EF 2 (Emotional Regulation/Shifting)											
1.	381.76	15	<.001	–	–	–	.091	.084, .099	0.069	.991	–
2.	435.20	18	<.001	57.16	3	<.001	.089	.092, .096	0.071	.991	0 Pass
3.	494.27	25	<.001	58.94	7	<.001	.080	.074, .086	0.072	.989	.002 Pass
4.	508.65	29	<.001	17.14	4	<.001	.075	.069, .081	0.072	.989	0 Pass
Prosocial Skills											
1.	374.82	29	<.001	–	–	–	.064	.058, .070	0.035	.984	–
2.	370.01	33	<.001	8.51	4	.075	.059	.054, .065	0.036	.984	0 Pass
3.	438.08	37	<.001	84.60	4	<.001	.061	.056, .066	0.037	.981	.003 Pass
4.	447.86	42	<.001	31.47	5	<.001	.057	.053, .062	0.038	.981	0 Pass

Note. RMSEA = Root Mean Square Error of Approximation; 90% CI: 90% Confidence Interval for RMSEA; WRMR = Weighted Root Mean Square Residual; CFI =Comparative Fit Index; Δ CFI = change in Comparative Fit Index; if change <=.01, the model "passed". 1 = configural invariance; 2 = weak invariance; 3 = strong invariance; 4 = unique invariance.

Table A13

Fit Indices for Auto-Regressive Cross-Lagged Models Using Accelerated Longitudinal Design

Model	*N*	χ^2	*df*	RMSEA	[90% CI]	WRMR	CFI
Math	2,525	1166.93***	121	.059	[.055, .062]	.049	.979
Reading	2,525	1285.88***	121	.062	[.059, .065]	.055	.979

Note. RMSEA = Root Mean Square Error of Approximation; 90% CI: 90% Confidence

Interval for RMSEA; WRMR = Weighted Root Mean Square Residual; CFI = Comparative Fit Index

***$p < 0.001$.

Table A14

Structural Model Generated Within Time Correlations for Total Sample (N = 2,525)

	EF 1	EF 2	Prosocial	EF 1	EF 2	Prosocial
	Fall Kindergarten			Spring Kindergarten		
EF 2	.78***	–	–	.89***	–	–
Prosocial	.79***	.72***	–	.74***	.76***	–
Math	.27***	.15***	. 18***	.22***	.30***	.19***
Reading	.22***	.12***	.11***	.16***	.11***	.06*
	Fall Grade 1			Spring Grade 1		
EF 2	.87***	–	–	.89***	–	–
Prosocial	.77***	.75***	–	.71***	.71***	–
Math	.13***	.07**	. 05	.18***	.15***	.07***
Reading	.21***	.17***	.08*	.18***	.16***	.09***
	Fall Grade 2			Spring Grade 2		
EF 2	.88***	–	–	.92***	–	–
Prosocial	.77***	.76***	–	.74***	.71***	–
Math	.19***	.15***	.12***	.07*	.04	.03
Reading	.12***	.10***	.06*	.10***	.09**	.08**

Note. Correlations reported for EF 1, EF 2, and Prosocial Skills were generated based on the Math model, however, these correlations are identical for the Reading model.

EF 1 = Executive Functioning 1 (Working Memory, Plan, Problem Solving);

EF 2 = Executive Functioning 2 (Shifting, Emotional Regulation);

Prosocial = Prosocial Skills.

**p*<0.05; ***p*<0.01; ****p*<0.001.

Table A15

Structural Models' Stability Coefficients (N's = 2,525)

	Model 1 Math	Model 2 Reading
Achievement (Math or Reading Models)		
Spring K *on* Fall K	.43(.02)***	.68(.02)***
Fall Gr. 1 *on* Spring K	.57(.02)***	.83(.01)***
Spring Gr. 1 *on* Fall Gr.1	.45(.02)***	.79(.01)***
Fall Gr.2 *on* Spring Gr.1	.48(.02)***	.85(.01)***
Spring Gr.2 *on* Fall Gr.2	.42(.03)***	.85(.01)***
Prosocial Skills		
Spring K *on* Fall K	.91(.02)***	.91(.02)***
Fall Gr. 1 *on* Spring K	.31(.03)***	.31(.03)***
Spring Gr. 1 *on* Fall Gr.1	.90(.01)***	.90(.01)***
Fall Gr.2 *on* Spring Gr.1	.34(.03)***	.34(.03)***
Spring Gr.2 *on* Fall Gr.2	.92(.02)***	.92(.02)***
Executive Functioning 1 (EF 1)		
Spring K *on* Fall K	.77(.02)***	.77(.01)***
Fall Gr. 1 *on* Spring K	.66(.02)***	.66(.02)***
Spring Gr. 1 *on* Fall Gr.1	.79(.02)***	.79(.02)***
Fall Gr.2 *on* Spring Gr.1	.64(.02)***	.64(.02)***
Spring Gr.2 *on* Fall Gr.2	.83(.01)***	.83(.01)***
Executive Functioning 2 (EF 2)		
Spring K *on* Fall K	.71(.02)***	.71(.02)***
Fall Gr. 1 *on* Spring K	.50(.02)***	.50(.02)***
Spring Gr. 1 *on* Fall Gr.1	.71(.01)***	.71(.01)***
Fall Gr.2 *on* Spring Gr.1	.52(.02)***	.52(.02)***
Spring Gr.2 *on* Fall Gr.2	.78(.01)***	.78(.01)***

Note. Achievement refers to either Math (second column) or Reading (third column).

Stability coefficients presented are the completely standardized parameter estimates.

EF 1 = Working Memory/Plan/Problem Solve; EF 2 = Shifting/Emotional Regulation

Fall K = Fall Kindergarten; Spring K = Spring Kindergarten; Gr. = Grade.

*p<0.05; **p<0.01; ***p<0.001.

Table A16

Structural Models' Parameter Estimates for the Cross-Lagged Paths (N's = 2,525)

	Model 1 (Math)	Model 2 (Read)
Executive Functioning 1 (EF 1)		
Spring K *on* Prosocial (Fall K)	-.02(.02)	-.02(.02)
Fall Gr. 1 *on* Prosocial (Spring K)	-.06(.03)*	-.06(.03)*
Spring Gr.1 *on* Prosocial (Fall Gr.1)	.03(.01)*	.03(.01)*
Fall Gr.2 *on* Prosocial (Spring Gr.1)	-.04(.03)	-.04(.03)
Spring Gr.2 *on* Prosocial (Fall Gr.2)	.03(.02)	.03(.02)
Executive Functioning 2 (EF 2)		
Spring K *on* Prosocial (Fall K)	.07(.02)***	.07(.02)***
Fall Gr. 1 *on* Prosocial (Spring K)	.04(.03)	.04(.03)
Spring Gr.1 *on* Prosocial (Fall Gr.1)	.12(.01)***	.12(.01)***
Fall Gr.2 *on* Prosocial (Spring Gr.1)	.05(.03)	.05(.03)
Spring Gr.2 *on* Prosocial (Fall Gr.2)	.07(.02)***	.07(.02)***
Prosocial Skills		
Spring K *on* EF 1 (Fall K)	-.11(.02)***	-.11(.02)***
on EF 2 (Fall K)	-.01(.02)	-.01(.02)
Fall Gr. 1 *on* EF 1 (Spring K)	.03(.03)	.03(.03)
on EF 2 (Spring K)	.24(.03)***	.24(.03)***
Spring Gr.1 *on* EF 1 (Fall Gr.1)	-.13(.02)***	-.13(.02)***
on EF 2 (Fall Gr.1)	.04(.02)*	.04(.02)*
Fall Gr.2 *on* EF 1 (Spring Gr.1)	-.11(.03)***	-.11(.03)***
on EF 2 (Spring Gr.1)	.30(.02)***	.30(.02)***
Spring Gr.2 *on* EF 1 (Fall Gr.2)	-.15(.02)***	-.15(.02)***
on EF 2 (Fall Gr.2)	.16(.02)***	.16(.02)***
Academic Achievement		
Spring K *on* EF 1 (Fall K)	.42(.05)***	.38(.04)***
on EF 2 (Fall K)	-.11(.04)**	-.16(.02)***
on Prosocial (Fall K)	-.01(.04)	-.07(.03)*
Fall Gr. 1 *on* EF 1 (Spring K)	.29(.05)***	.18(.03)***
on EF 2 (Spring K)	-.02(.05)	-.04(.02)
on Prosocial (Spring K)	-.06(.04)	-.06(.02)*
Spring Gr. 1 *on* EF 1 (Fall Gr.1)	.24(.04)***	.21(.02)***
on EF 2 (Fall Gr.1)	-.01(.04)	-.06(.02)**
on Prosocial (Fall Gr.1)	-05(.03)	-.03(.02)
Fall Gr.2 *on* EF 1 (Spring Gr.1)	.32(.05)***	.13(.03)***
on EF 2 (Spring Gr.1)	-.13(.06)*	-.06(.03)
on Prosocial (Spring Gr.1)	-.06(.04)	-.03(.02)
Spring Gr.2 *on* EF 1 (Fall Gr.2)	.31(.05)***	.06(.03)
on EF 2 (Fall Gr.2)	-.12(.06)*	.01(.02)
on Prosocial (Fall Gr.2)	-.02(.04)	-.05(.02)

Note. Parameter estimates presented are completely standardized. *p<0.05; **p<0.01; ***p<0.001.

APPENDIX B

FIGURES

Figure B1

Proposed Relationship Among Executive Functioning, Prosocial Skills, and Academic Achievement Over 4 Waves

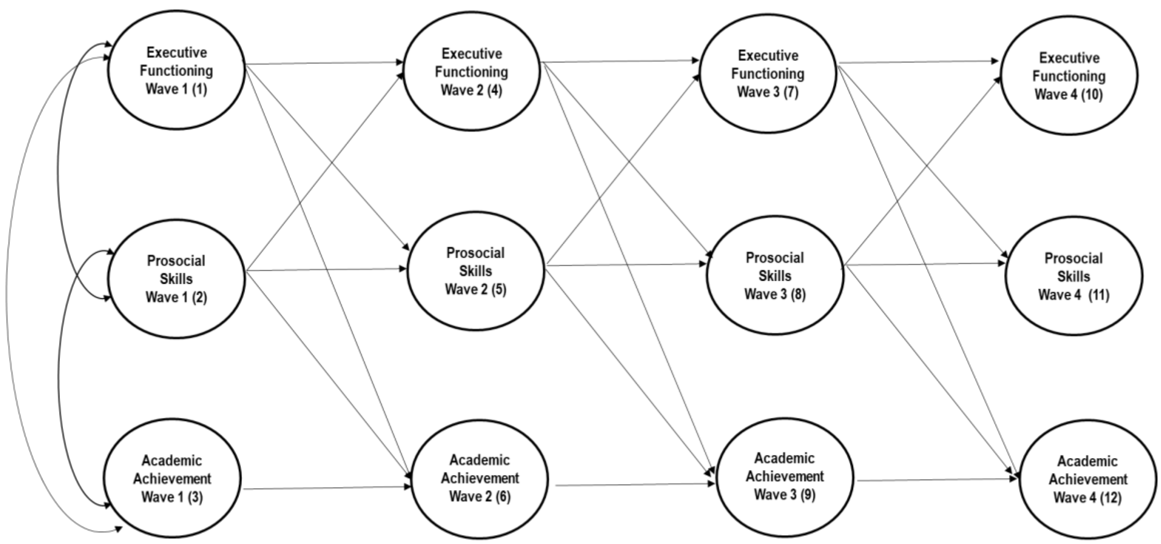
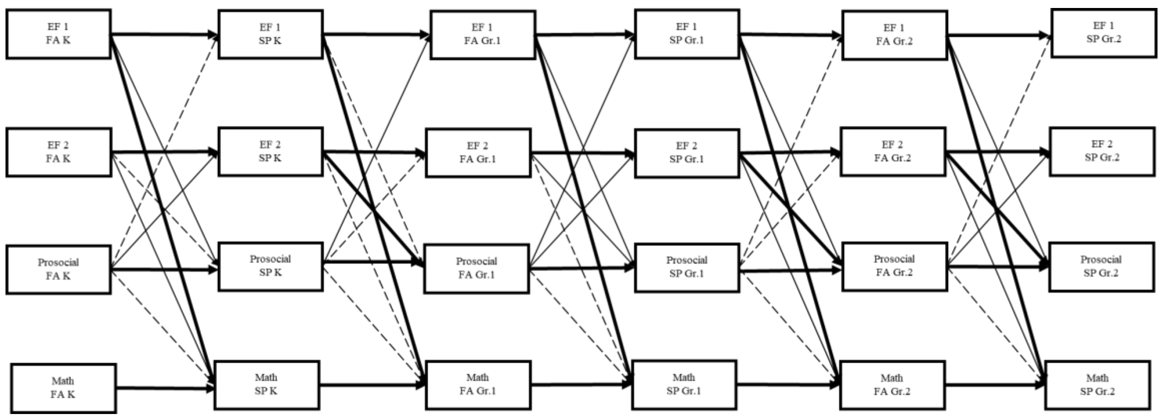


Figure B2

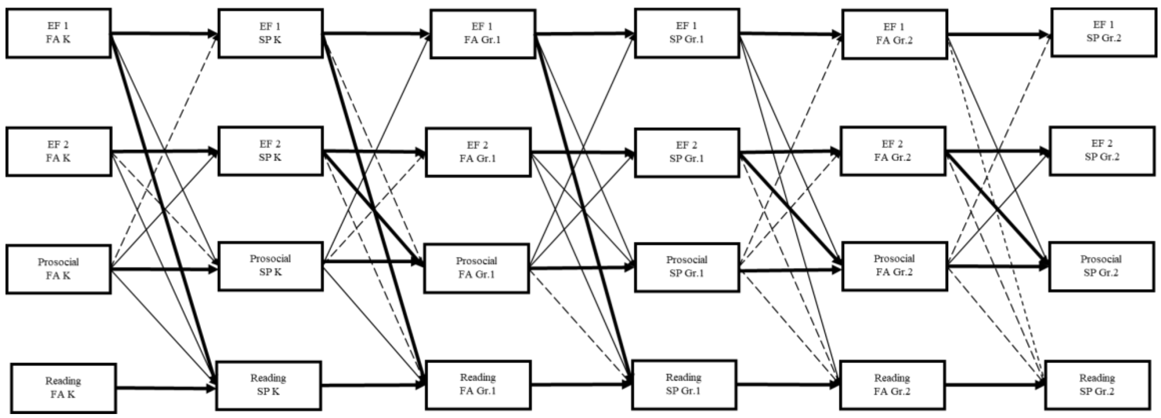
Model 1 (Math) Auto-Regressive Cross-Lagged Model for Kindergarten through Grade 2 Using Accelerated Longitudinal Design



Note. Significant paths are represented by solid lines. Significant and meaningful pathways are represented by heavy set lines. Nonsignificant paths are represented by dashed lines. Effect sizes are omitted for ease of presentation.

Figure B3

Model 2 (Reading) Auto-Regressive Cross-Lagged Model for Kindergarten through Grade 2 Using Accelerated Longitudinal Design



Note. Significant paths are represented by solid lines. Significant and meaningful pathways are represented by heavy set lines. Nonsignificant paths are represented by dashed lines. Effect sizes are omitted for ease of presentation.

APPENDIX C

COMMONLY CITED MEASURES OF EXECUTIVE FUNCTIONING AND SOCIAL EMOTIONAL LEARNING SKILLS

Measures of Executive Functioning

Executive functioning tests. EF tests are laboratory measures that ask the respondent to follow a task's instructions to measure response time and/or accuracy (Morrison & Grammar, 2016). The purpose of this section herein is twofold. First, this section functions to familiarize the reader with EF tests by reviewing those that are among the most commonly cited in the literature. This review is organized by component of EF. Second, this section functions to review methodological issues previously identified with EF tests.

Review of executive functioning tests by component. Every effort has been made in this section to review the most commonly used EF tests, based on Jacob and Parkinson's (2015) identification. In limiting this review of EF tests to the most commonly cited tests, it is hoped that not only will readers be better equipped to deepen their understanding of the components of EF, but in turn to deepen their familiarity with EF tests more generally. In the hopes that familiarity breeds comfort (not contempt), the reader may be better equipped to evaluate studies using various EF tests in the literature.

Working memory. The Wechsler scales are the most widely used measure of intelligence worldwide (Flanagan & Kaufman, 2009), with one of its scales, the Wechsler Intelligence Scale for Children- Fourth Edition (WISC-IV; Wechsler, 2003) containing the component of working memory. WISC – IV subtests include the Digits Span Backward and Letter-Number Sequencing, both with good test-retest reliability (rs = .83; Alloway et al., 2008). Recent meta-analyses suggest the Digit Span Backward (Baggetta & Alexander, 2016; Jacob & Parkinson, 2015) and Backward Word Span tests (Jacob & Parkinson, 2015) are the most frequently used test for working memory. These tests involve an experimenter reading a sequence of numbers/digits, words, or letter-numbers before the child is asked to repeat the sequence in backwards order (e.g., see Alloway et al., 2008). Validity issues of backward word-span type tests include not discriminate between various types of working memory (Kane et al., 2004), questionable factor structures (Maricle & Avirett, 2012), and capturing a related but distinct construct, that of executive attention (Conway et al., 2005). Other types of working memory span tests (e.g., counting, operations, and reading) are also widely used and tend to have adequate reliability (Alloway et al., 2008; Conway et al., 2005) and good convergent, discriminant, and predictive validity (Conway et al., 2005).

Inhibition. In Baggetta and Alexander's (2016) meta-analysis (N = 106), inhibition was assessed 92 times by 28 different EF tests. While there are many ways to assess inhibition, the Stroop Color-Word Interference Test (SCWT; Stroop, 1935) remains the most frequently used EF test (Baggetta & Alexander, 2016; Jacob & Parkinson, 2015). There are multiple variations of the SCWT, translated into multiple languages (see Homack & Riccio, 2004). Developed for attention and cognitive flexibility, the test is primarily used to measure inhibition of thoughts (versus behaviors). The SCWT involves color-word processing such that it juxtaposes automatic process (word reading) against a controlled process (color naming) (MacLeod, 1992). For example, the respondent is asked to read the word "yellow" written in blue ink. The Color-Word task involves creating a composite score based on three separate scores obtained by identifying the number of words (Word task), number of bar colors (Color), and number of color words (Color–Word) within a certain amount of time (Homack & Riccio, 2004).

The psychological processes involved to master the test are referred to as the Stroop effect and have been the subject of hundreds of studies and is judged to have good reliability and validity (for a review, see MacLeod, 1991). One meta-analysis conducted using the SCWT and its variations in children with clinical disorders ($K = 33$) suggests the test has good sensitivity and specificity for certain populations with EF deficits (Homack & Riccio, 2004). The task has been criticized for task impurity (Miyake & Friedman, 2012), such that it is influenced by reading skills (Golden, Espe-Pfeifer, & Wachsler-Felder, 2000; MacLeod, 1991). As such, opponents argue the test should not be used to assess EF (Reynolds & MacNeill, 2014). Archibald and Kerns' (1999) proposed solution is the Fruit Stroop Task (test-retest reliability $r = .87$-$.93$), which asks respondents to name the color of rectangles, and properly and improperly colored fruits and vegetables.

Shifting. In Baggetta and Alexander's (2016) meta-analysis ($N = 106$), shifting was assessed 80 times by 33 different EF tests. The Wisconsin Card-Sorting Test (WCST; Berg, 1948) is among the most common measure, with a revised test-manual published thereafter (Heaton, 1981). Respondents are asked to determine the rules needed such that cards (two decks of 32 cards each) are sorted in the order of color (i.e., red, green, yellow, blue), shape (i.e., triangle, star, cross, circle), and then number (i.e., 1-4) (Goldstein & Naglieri, 2014; Jacob & Parkinson, 2015). Scoring is based on the time it takes and the number of errors made when trying to figure out the sorting rule, which changes throughout the test, a common scoring method among shifting tests (Monsell, 2003). Variations include an adapted version of the Delis-Kaplan Executive Function

System (D-KEFS; Delis, Kaplan, & Kramer, 2001), a common battery of EF (Baggetta & Alexander, 2016).

There is a dearth of studies on the reliability and validity of the WCST in normally developing children, with data on adults (e.g., Bowden et al., 1998) suggesting adequate interrater reliability. While age effects are reported to be small (Lezak, 2004), data from the WCST administered in various countries suggests performance on the test improves as children age (e.g., see Yeniceri & Altan-Atalay, 2011). For example, in Bujoreanu and Willis' (2008) study of American children (N = 196), while error rates were constant across the age groups (6, 11, and 18 years), the youngest age group had the poorest performance, suggesting that number sorting improves with age. One meta-analysis suggested the test has good sensitivity and specificity for ADHD populations (Romine et al., 2004).

Planning/Problem Solving. Two common EF tests of planning in school-aged children are the Tower of Hanoi (ToH; Piaget, 1974/1976) and the Tower of London (ToL; Shallice, 1982) with graduated levels of difficulty to assess developmental differences. Both these tests are disk-transfer tasks involving moving objects onto pegs according to a set of rules and within a certain number of moves (Welsh & Huisinga, 2001). The ToH uses disks of varying sizes and three same-sized pegs and the ToL, which is based on the ToH, uses colored balls and three different-sized vertical pegs. Children's rating scales (e.g., BRIEF; BDEFS-CA; CHEXI) also measure planning (refer to subsection on *Rating Scales*).

There are a variety of variations of the ToH and ToL tests, with few reviews or meta-analyses conducted, making it difficult to report on their reliability and validity in a consistent manner. In order to address issues of low reliability for the ToH, proposed revisions have been suggested (e.g., see Ahhoniska et al, 2000; Welsh and Huisinga, 2001). A review of studies on the reliability for the ToL suggest a range of poor to excellent (e.g., test-retest, split-half, intra-class correlation), with proposed revisions including increasing the number of problems a respondent is asked to complete (Tunstall et al., 2014). With regards to construct validity, Miyake et al. (2000) suggest inhibition and planning load on the same factor, however, this analysis was conducted on a sample of adults (*N* = 137). Bishop, Aamodt-Leeper, Creswell, McGurk, and Skuse (2001) suggest ToH, administered to children aged 7 to 15 (*N* = 238), was not significantly correlated with inhibition. Similarly, for the ToL, results of a factor analysis conducted by Bull, Epsy, & Senn (2004) on a sample (*N* = 118) of normally developing three-to six-year-olds suggested the ToL did not load on planning but loaded on inhibition. Perhaps this finding is due to the mix of preschoolers and elementary school aged children in the sample. For example, Baughman and Cooper (2007) found that children (*N* = 34) who had fewer occurrences of rule breaking (execution) on multiple tasks using a computational approach to the ToL were older (aged 5 and 6) while younger children (aged 3 and 4) had higher occurrences of rule breaking, a phenomena attributed to the development of inhibition among the older children. In short, many childhood EF researchers use the ToH or ToL as simultaneous measures of inhibition and planning (e.g., refer to studies in section Developmental Trajectories). With regards to addressing construct validity, Zelazo and Muller (2010) suggest planning tests distinguish between planning and later phases such as execution. Both tests have good specificity in identifying prefrontal damage related to planning in various clinical populations (Welsh & Huisinga, 2001).

Emotional regulation. Few EF tests independently examine emotional regulation, with the Children's Gambling Task (Kerr & Zelazo, 2004) being one such test (Peterson & Welsh, 2014). Developed for preschoolers, it is suitable and often used to assess school-aged children (Zelazo & Muller, 2010), with developmental differences easily observed (Kerr & Zelazo, 2004). This test has two distinguishable decks of cards that have varying amounts of happy and sad faces on each card and children loose or win candy based on the card selected (Zelazo & Muller, 2010). The ratio of rewards to losses in each deck if different with one deck having a greater and more variable reward to risk ratio than the other deck. Because the first half of cards drawn do not count towards the total score, the child has a chance to notice the distinction between the more and less risky deck of cards and thus the proportion of disadvantageous choices is evident. At the end of fifty card draws, the pieces of candy remaining (total score) represents emotionally significant consequences of decision making.

Methodological problems with executive functioning tests. There are a host of problems associated with EF tests, notably, with respect to validity. For example, one of the largest issues is the task impurity problem; some test may be used to measure one component of EF, but simultaneously measure other non-EF cognitive functions (Miyake & Friedman, 2012). The problem remains despite attempts to rectify it, such as Miyake et al.'s (2000) latent variable approach that tease out variance attributed to non-EF factors (Miyake & Friedman, 2012). Along these lines, researchers do not consistently attribute a specific EF tests to the same component (Best & Miller, 2010; Lezak, 2004; Morris, 1996), possibly because most EF tests were developed to assess other psychological functioning (Lezak, 2004), unlike EF rating scales. For example, the Wisconsin Card Sorting Task and the children's version (Dimensional Change Card Sorting Task; Frye, Zelazo & Palfai, 1995) have been described as inhibition tasks by some and shifting tasks by others (Garon et al., 2008). Furthermore, a review by Mahone et al. (2002) suggests some neuropsychological tests of EF are moderately correlated with measures of IQ (e.g., WISK-R & WISC III), contrary to EF rating scales (Alderman et al., 2003; Barkley & Murphy, 2011). In addition, a plethora of tests make comparison across research findings a challenge. In a review of 106 studies measuring EF, 109 different tests were identified with 27% of these tests contained two to six components, in addition to sometimes being used to measure EF as a global construct (Baggetta & Alexander, 2016). Most studies (54%) only use one task (versus multiple tasks) to measure a component, raising issue of validity and reliability. Furthermore, some studies (e.g., Huizinga et al., 2006) note EF tests of the same component (e.g., inhibition) on similar populations do not yield the same results.

EF tests emerged out of research derived from frontal-lobe functioning in adults whereas rating scales emerged from theories using a developmental framework (Brocki & Bohlin, 2004). Consequently, EF tests are less adept at capturing deficit in everyday functioning (Buchanan et al, 2010; Gioia, Isquith, & Guy, 2001) over time (Barkley & Murphy, 2011). For example, many EF tests (e.g., Trail Making Test, Stroop, Wisconsin Card Sorting) account for a much smaller amount of variance in everyday functioning compared to rating scales (e.g., DEX, BAFT, DEFS/DEFI) (Chaytor, Schmitter-Edgecombe, & Burr, 2006; Barkley & Murphy, 2011, Barkley & Fischer, 2011) because:

(a) these tests observe the subject over small windows of time (e.g., 5-30 minutes) compared to rating scales (Barkley, 2012b); and (b) a subject taking an EF test is unlikely to repeat this activity in her every day functioning (Alderman, Burgess, Knight, & Henman, 2003). Barkley (2012a, 2012b) further argues that EF tests, unlike rating scales, do not measure social behaviors and skills associated with EF over time (i.e., hours, weeks, months) such as social interdependence (e.g., cooperation), daily social exchanges (e.g., sharing, turn-taking, reciprocity), and mutualism (e.g., reciprocal concerns for the long-term welfare of others) for the purpose of obtaining future goals. In other words, EF tests and rating scales may capture different aspects of EF (Chaytor, Schmitter-Edgecombe, & Burr, 2006) with Toplak, West, and Stanovich (2013) arguing that EF tests capture efficiency of cognitive abilities while rating scales capture success in goal pursuit.

Efforts to overcome measurement issues associated with EF tests are underway, such as the proposed revised version to the National Institutes of Health Toolbox Cognitive Battery® (Weintraub et al., 2013; Zelazo et al., 2013). This measure of EF and attention was developed in conjunction with a large panel of experts based on EF tests commonly associated with working memory (NIH's List Sorting Working Memory Test), shifting (Dimensional Change Card Sort; Zelazo, 2006), and inhibition (Eriksen Flanker Task; Eriksen & Eriksen, 1974) to develop a measure of EF over the life course (for a study on the factor validity, see Mungas et al., 2013).

To conclude, I would like to draw the reader's attention to an addition concern I have, not with EF tests, but the way EF tests are used in the EF literature. A plethora of EF tests are used in EF research but the test names are often not referenced in the EF literature, including book chapters, meta-analyses, meta-reviews, and literature review sections in article. In other words, when EF components are discussed in terms of trajectories or their relation to other variables (e.g., achievement) the research reviewed is not differentiated by EF test. In an effort to address this issue, the various EF tests used in studies described herein are explicitly named, cited, and referenced. Refer to the second column of Table A1 for a concise overview of additional EF tests not reviewed in this section above that are cited in later sections of this literature review.

Rating scales. Rating scales involve a series of questions that measure behavioral manifestations of one or more components of EF via self or other (teacher/parent) reports (Jacob & Parkinson, 2015). Three prominent children's rating scales are discussed below. Table A1 (column 3) includes a description of these three scale items to better convey each component of EF in the context of every day functioning.

Behavior Rating Inventory of Executive Functioning (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000). The BRIEF is one of the two most commonly used rating scales to assess EF in children (Catale et al., 2015) and was developed by a team of neuropsychologists. BRIEF is an 86 item, 3-point Likert scale with two versions (parent and teacher) to rate children aged 5 to 18, with different age groups scored separately. The measures were developed and standardized on large samples ($n_{Parent\text{-}Report}$ = 1419 and $n_{Teacher\text{-}Report}$ = 719) of ethnically diverse American children, suggesting a two-factor solution accounting for 74-83% of the variance of EF (Gioia et al, 2000, Gioia, Isquith, & Guy, 2001): (a) *Metacognition Index* (MCI) measures problem-solving abilities around self-managing tasks and monitoring one's performance (subscales: initiate, working memory, plan/organize, monitor, and organization of materials); and (b) *Behavior*

Regulation Index (BRI) measures the child's ability to use inhibitory control by modulating emotions/responses and behavior and by using shift cognitive set (subscales: inhibit, shifting, and emotional control) (Isquith & Gioia, 2013).

With over 400 peer-reviewed studies (Roth, Erdodi, McCullah, & Isquith, 2015), the scale's merits include: good reliability among normative and clinical samples for both versions (α = .80-.98) (Gioia et al., 2000); good reliability in translated versions (Fournet et al., 2015); good convergent and discriminant validity (Gioia et al., 2001; Hoefling, 2016); good specificity with ADHD populations (Gioia et al., 2000); and good sensitivity by not overly identifying EF deficits (Roth et al., 2015). However, there is some controversy on the specified two factor solution (e.g., Egeland & Fallmyr, 2010; Fournet et al., 2015).

Childhood Executive Functioning Inventory (CHEXI; Thorell & Nyberg, 2008). The second of the two most commonly used rating scales to assess EF in children (Catale et al., 2015) is the CHEXI, a 26-item, 4-point Likert scale with two versions, one for parents and one for teacher (Thorell & Nyberg, 2008). CHEXI was initially tested on Swedish children (N = 130, M_{age} = 6 yrs.) with results suggesting a two-factor solution (for both versions), supported by additional testing on Belgian children aged 8 to 11 (N = 242) by Catale et al. (2015) [$\chi^2_{(251)}$ = 520,336, p < .001, RMSEA .07, CFI .97, and SRMR .06]. The two factors are labeled working memory/planning and inhibition/ regulation, accounting for 41% of the variance of EF using the parent version and 67% of the variance using the teacher version (Thorell & Nyberg, 2008; Thorell, Veleiro, Siu, & Mohammadi, 2013). The working memory/planning factor (13 items) represents having difficulties such as remembering and understanding instructions, getting stuck, multitasking, and planning (e.g., remembering things, executing multi-step tasks) (Catale et al., 2015). The inhibition/regulation factor (11 items) represents the cognitive and motivational aspects of inhibitory control such as being overly excited, acts before thinking, following through, and concentration. As such, this factor measures behaviors, responses, and thoughts involved in inhibition. The working memory factor is partially derived from Baddeley and Hitch's (1974) dual component model of working memory, which functions to store and process verbal and spatial information. The inhibition/regulation factor is derived from Barkley's (1997) hybrid model in which these factors, in addition to working memory, are seen as constituting the major EF deficits in children with ADHD. The CHEXI had high specificity and sensitivity to distinguish between ADHD and controls in Thorell, Eninger, Brocki, and Bohlin's (2010) study of five year-olds ($n_{\text{ADHD}} = 15$; $n_{\text{control}} = 30$) and in Catale et al.'s (2015) studies of Belgian ($n_{\text{ADHD}} = 25$; $n_{\text{control}} = 25$) and Swedish ($n_{\text{ADHD}} = 62$; $n_{\text{control}} = 62$) children aged 8 to 11.

Barkley Deficits in Executive Functioning Scale – Children and Adolescents (BDEFS – CA; Barkley, 2012b). This scale is a 70-item, 4-point parent-report scale that is based on the Barkley Deficits in Executive Functioning Scale (BDEFS for Adults; Barkley, 2012c). The BDEFS – CA is rooted in Barkley's research on ADHD (e.g., Barkley, Murphy, & Fischer, 2008) and tested on a normative sample of American children ages 6 to 17 years ($N = 1{,}922$). Barkley's (2012a; 2012b) assumes that EF is highly important for social functioning and thus the BDEFS includes emotional regulation, similar to the BREIF's second factor (BRI) of modulating/ controlling one's emotional response to the situation or stressor (Gioia at al., 2000). Another distinction of the BDEFS centers on Barkley's (2012b) definition of EF: "acts of self-regulation across time towards future goals" (p. 8). These acts, or self-directed everyday life activities, are represented by five subscales: a) self-regulation of emotions; b) self-organization/problem solving; c) self-management of time; d) self-motivation; e) self-restraint (inhibition of e.g., behaviors, responses, and thoughts). Barkley (2012b) reports these five subscales account for 68.6% of the variance in EF as identified by a principal component factor analysis using both varimax and promax solutions. The five factors are moderately to highly intercorrelated (r = .64 - .89, p < .05), with the shared variance thus accounting for between 39% and 70%, Barkley (2012b) argues the construct is unitary in nature, represented by self-regulation. BDEFS – CA (parent version) has good test-retest reliability three to five weeks after initial completion (N = 86, r = .73 - .82), on par with Gioia et al.'s (2000) BRIEF (r = .76-.85), and Thorell and Nyberg's (2008) reporting of the CHEXI (r =.89). Barkley (2012b) also reports the BDEFS – CA has good construct, convergent, divergent, and criterion validity.

Measures of Social Emotional Learning Skills – Prosocial Behavior

While there is a plethora of measures of prosocial behavior, four of the most common measures are discussed, namely: (a) Devereux Student Strengths Assessment – Second Step® Edition (DESSA – SSE; Devereux Center for Resilient Children, 2012); (b) Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997); (c) the Social Skills Improvement System Rating Scales (SSIS – RS; Gresham & Elliott, 2008); and (d) the Social Skills Improvement System SEL Edition Rating Form (SSIS – SEL RF; Gresham & Elliott, 2017).

Devereux Student Strengths Assessment – Second Step® Edition (DESSA – SSE; Devereux Center for Resilient Children, 2012). DESSA-SSE is a rating scale for teachers to report socio-emotional competencies of their students in Kindergarten through 8th Grade that they observed over the last month based on 36 items rated on a 5-point Likert scale (0 = never to 4 = very frequently). These socio-emotional competencies function to assess protective factors, which Naglieri and LeBuffe (2006) define as "factors that lead to resilient outcomes" (p.109), based on four subscales (9-items each) in addition to a social-emotional composite subscale (36-items): skills for learning (e.g., does routine tasks or chores without being reminded); empathy (e.g., offers to help somebody); emotion management (e.g., wait his/her turn), and problem solving (e.g., take steps to achieve goals). The same version of the scale is administered to students in each grade.

While there is scant information on the validity and reliability of the DESSA-SSE, this survey is based on DESSA (LeBuffe, Shapiro, & Naglieri, 2009), a rating scale for parents and teachers to report on resiliency behaviors of children from Kindergarten through 8th Grade based on eight dimensions of resiliency consisting of 72 items rated on a 5-point Likert scale (Nickerson & Fishman, 2009). LeBuffe, Shapiro, & Naglieri (2009) developed the scale based on identifying 765 characteristics of resilient children. A study of largely Caucasian parents (*n* = 133) and teachers (*n* = 94) by Nickerson and Fishman (2009) suggests the DESSA (LeBuffe, Shapiro, & Naglieri, 2009) has good reliability (α = .87-.93) and that it is a valid measure of resilience. Validity was assessed via convergent validity, whereby moderate to high positive correlations (positive and negative, respectively) were reported on both the parent and teacher rating scales using the Emotional Rating Scales (BERS–2; Epstein, 2004) and the Behavior Assessment System for Children–2 Adaptive Skills composite (BASC–2, Reynolds & Kamphaus, 2004). Divergent validity was suggested using subscales of externalizing problems, internalizing problems, and adaptive skills on the BASC–2 (Behavioral Symptom Index clinical version; Reynolds & Kamphaus, 2004), based on moderate to high negative correlations for parent and teacher rating scales.

Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997). The SDQ is comprised of 25-items rated by teachers or parents (two versions) on a 3-point scale to assess the strengths and difficulties of children aged 4-16 years. Children's difficulties are assessed via four subscales (i.e., conduct problems, emotional symptoms, hyperactivity, and peer relationship problems), each with five items. Children's strengths are assessed via a fifth subscale (prosocial behavior), also with five items. The items on the prosocial scale are as follows: "considerate of other people's feelings"; "shares readily with other children"; "helpful if someone is hurt, upset, or feeling ill"; "kind to younger children"; and "often offers to help others". The same version of the scale is administered to students in each grade. The SDQ is a widely used measure for screening and research purposes that has been translated into over 60 languages (Stone et al., 2010).

A meta-analysis of 47 studies examined the psychometric properties of the parent and teacher versions in children aged 4-12 (N = 131,223). Reliability for the parent and teacher versions are satisfactory. Specifically, while the internal consistency weighted mean (k = 26) on the parent version is borderline acceptable (α = .67), the mean on the teacher version is good (α =.82). The test-re-test weighted mean (k = 8) on the parent version is borderline acceptable (r = .65), and on the teacher version is acceptable (r = .79). The inter-rater reliability between parent and teacher versions are satisfactory (k = 6, r = .26), according to standards suggested by Achenbach, McConaughy, Howell, and

Masters (1987). With respect to Stone et al.'s (2010) evaluation of construct validity, when examining factor loadings on the five prosocial scale items across studies, more studies had factor loadings of .70 or above compared to ranging between .40-.69 for the teacher version (N = 19,105, k = 7) than for the parent version (N = 43,274, k = 14). Further validity of the SDQ is suggested by high predictive and concurrent validity (Goodman, 1997), as well as criterion and discriminant validity (Goodman & Scott, 1999).

Social Skills Improvement System Rating Scales (SSIS – RS; Gresham & Elliott, 2008). Gresham and Elliott (2008) developed the Social Skills Improvement System-Rating Scales (SSIS-RS), a multi-informant (teachers, parents, students), four-point frequency scale used to identify social skills deficits in children aged 3 to 18 years. Derived from the Social Skills Rating System (SSRS; Gresham & Elliott, 1990), the SSIS-RS includes a social skills domain with seven subscales: assertion, cooperation, empathy, responsibility, self-control, communication, and engagement. SSIS – RS' social skills domain measures prosocial behavior (Gresham, Elliott, Vance, & Cook, 2011; Kettler, Elliott, Davies, & Griffin, 2011). SSIS – RS (Gresham & Elliott, 2008) and its parent scale the SSRS (Gresham & Elliott, 1990) are widely used. For instance, a meta-analysis of 85 different measures of social functioning from 1988 to 2010 found they were the most commonly cited rating scale of children's social skills in the research (Crowe, Beauchamp, Catroppa, & Anderson, 2011). Furthermore, these scales are the most widely used rating scales of social behaviors in the USA and some foreign schools (Gresham et al., 2011); and have been translated into a dozen different languages (e.g., SSIS – RS; Sherbow et al., 2015).

Social Skills Improvement System SEL Edition Rating Form (SSIS SEL – RF; Gresham & Elliott, 2017). All items from the social skills subscale and three items from the problem behavior subscale of the SSIS – RS (Gresham & Elliott, 2008) constitute a newly created scale: the Social Skills Improvement System SEL Edition Rating Form (SSIS SEL – RF; Gresham & Elliott, 2017). The SSIS SEL – RF was created to fill a gap in the SEL measures literature for a rating scale that aligns with the CASEL Core-5 (Elliott, Davies, Frey, Gresham, & Cooper, 2017; Gresham et al., in press). Face validity of the SSIS SEL – RF was assessed by having a panel of seven experts (i.e., SEL experts and teachers) assign each of SSIS – RS's social skills domain items to one of the CASEL Core-5 competencies until 100% consensus was reached. SSIS SEL – RF is a four-point rating scale for teachers, parents, and students (51, 51, and 46 items, respectively) with five subscales, or "Core Skills", one associated with each of the CASEL Core-5 competencies (self-awareness, self-management, social awareness, relationship skills, and responsible decision making). A sixth factor, academic competence (7-items), was created for the teacher version.

Reliability and validity of the SSIS SEL – RF (Gresham & Elliott, 2017) was assessed using nationally representative samples of teachers (n = 160), parents (n = 140), and students/self-reports (n = 224) reporting on American elementary and secondary school students. Reliability estimates of the SSIS SEL – RF for reports of elementary school children (aged 5-12) on the parent, teacher, and student surveys are excellent (α = .72-.97). To further assess the validity of the SSIS SEL – RF, three confirmatory factor analyses were performed, one for each version of the rating scale (parent, teacher, and, student) based on the data obtained to norm the original SSIS – RS on a sample of children (N = 4,700) aged 3 to 18 (Gresham et al., 2018). The values obtained using the RMSEA (0.062, 0.080, and 0.054, respectively) suggest the model is a fair fit to the data, however, the null hypothesis that the model was a good fit to the data was rejected based on the chi square statistic (ps < 0.05). Thus, there is preliminary evidence to suggest the SSIS SEL – RF is an adequate fit to CASEL's Core-5 model. Additional evidence of validity of the SSIS SEL – RF relates to convergent validity such that the parent, teacher, and student versions of its Core Skills subscales are moderately to highly correlated with the respective SEL – RS social skills domains (Gresham et al., 2018).

CASEL's Core-5, Prosocial Behavior, and the SSIS SEL – RF (Gresham & Elliott, 2017). As mentioned above, the social skills domain of the SSIS – RS (Gresham & Elliott, 2008), from which the SSIS SEL – RF (Gresham & Elliott, 2017) is derived, could be conceptualized as a measure of prosocial behavior (Gresham, Elliott, Vance, & Cook, 2011; Kettler, Elliott, Davies, & Griffin, 2011) that integrates the CASEL (2012) Core- 5 competencies described earlier (i.e., self-awareness, self-management, social awareness, relationship skills, responsible decision making).

Self-Awareness of thoughts and emotions as these relate to behavior within a social context is represented by eight items, including interacting and behaving well, such as "has difficulty waiting for turn" (#51). *Self-Management* of thoughts, emotions, and behaviors in regulating stress and impulses and sustaining motivation and goals is represented by 15 items, including turn-taking, compromising, self-responsibility, and "acts without thinking" (#47). *Social Awareness* (e.g., empathy, following norms, social support awareness) is represented by six items, including comforting others and classroom participation. *Relationship Skills* (e.g., communication, negotiation, cooperation, resisting social pressure, and proving/seeking social support) is represented by 14 items, including responding to and resolving disagreements, asking for help, acting responsibly, being kind, and bullying behaviors. Last, *Responsible Decision Making* for oneself and with regards to others is represented by eight items, including expression of a problem, staying calm, and expressing feelings.

APPENDIX D

COPYRIGHT PERMISSION

Creating Connections. Changing Lives. PAR®

Sent Via Email: ddesfoss@asu.edu

April 14, 2021

Danielle Desfosses, PhD
Arizona State University
PO Box 873701
Tempe, AZ 85287-3701

Dear Dr. Desfosses:

In response to your recent request, permission is hereby granted to you to include items 17, 27 and 60 from the Behavior Rating Inventory of Executive Function, Second Edition (BRIEF2) Teacher Form as an example in the appendix of your dissertation titled, *Cascade Model of Executive Functioning, Prosocial Skills, and Academic Achievement.* If additional material or further publication (i.e., Journal) is needed, then further permission from PAR is required.

This Agreement is subject to the following restrictions:

1) Any and all materials used will contain the following credit line:

   "Reproduced by special permission of the Publisher, Psychological Assessment Resources, Inc. (PAR), 16204 North Florida Avenue, Lutz, Florida 33549, from the Behavior Rating Inventory of Executive Function, Second Edition by Gerard A. Gioia, PhD, Peter K. Isquith, PhD, Steven C. Guy, PhD, and Lauren Kenworthy, PhD, Copyright 1996, 1998, 2000, 2015 by PAR. Further reproduction is prohibited without permission from PAR."

2) None of the material may be sold, given away, or used for purposes other than those described above.

3) Payment of a permission fee will be waived.

4) One copy of any of the material reproduced will be sent to PAR to indicate that the proper credit line has been used.

5) TWO COPIES of this Permission Agreement should be signed and returned to me to indicate your agreement with the above restrictions. I will then sign it for PAR and return a fully executed copy to you for your records.

Sincerely,

Vicki M. McFadden

Permissions Specialist

vmark@parinc.com

1-800-331-8378 (phone) 1-800-727-9329 (fax)

ACCEPTED AND AGREED:

BY:

DANIELLE DESFOSSES, PhD

DATE: April 14, 2021

PAR Customer No.: 237841

BY:

VICKI M. MCFADDEN

DATE: April 15, 2021

BRIEF2 Desfosses Teacher Sample Items only in dissertation (Arizona State Univ) - 4-14-2021

Subject: FW: Contact Us Form Submission - Permission to use DESSA-SSE
Date: Wednesday, March 17, 2021 at 1:02:56 PM Pacific Daylight Time
From: Jennifer Robitaille
To: danielle.desfosses@asu.edu
Attachments: image001.png, image002.png, image003.png, image004.png, image005.png, image006.png, image007.png

Hi Danielle,

Thank you for your interest in the DESSA-SSE! Please consider this email response as permission to reproduce the item content as you described in your email below in both your dissertation and subsequent journal article as applicable. We ask that you do not use this information for purposes other than that which is stated below and that you appropriately reference and acknowledge the DESSA-SSE in your manuscript(s).
If you have any questions please feel free to reach out to me directly. I look forward to seeing and learning from your findings!

Best,
Jennifer
Jennifer Robitaille, M.S., PMP
Director of Research and Development
Aperture Education
570-490-4397

From: Rudee Abello <Rudee.Abello@devereux.org>
Sent: Wednesday, March 10, 2021 5:50 PM
To: Jennifer Robitaille <JRobitaille@Apertureed.com>; Alyssa Ciarlante <ACiarlante@Apertureed.com> **Cc:** Susan Damico <sdamico@devereux.org>
Subject: FW: Contact Us Form Submission

Hi, Jen and Alyssa:

Hope you're both doing well!

I am wondering if you could respond to the inquiry below? She is requesting a response no later than 3/23/2021. Let me know.
Thank you!

Rudee Abello | Marketing Specialist Pronouns: she / her / hers
Devereux Center for Resilient Children
444 Devereux Dr.
Villanova, PA 19085
(w) 610-542-3189

From: DCRC Website <dcrc@devereux.org> **Sent:** Wednesday, March 10, 2021 5:38 PM
To: PA ED DECA <deca@devereux.org> **Subject:** Contact Us Form Submission
Name: Danielle Desfosses
Email: Danielle.desfosses@asu.edu
Country: United States
State: AZ
Agency: Grad student on behalf of Arizona State University

I'm a doctoral student at Arizona State University and in preparing to submit my dissertation for publication(s), it has just been brought to my attention that I need to request your permission to reproduce a small percentage of content from your rating scale the Devereux Student Strengths Assessment – Second Step® Edition (DESSA – SSE; Devereux Center for Resilient Children, 2012). Specifically, this content refers to 14 items. In no way does the inclusion of these items enable readers to reproduce any of the scale's four subscales [(skills for learning (5 items - #1, 6, 7, 12, 23), empathy (0 items), emotional management (5 items: #21, 25, 28, 31, 36), problem-solving (4 items: #4, 11, 30, 33)]. I obtained these items from a secondary dataset that belongs to my dissertation advisor, Dr. Sabina Low, who purchased the DESSA – SSE.

Including these items is pivotal to my dissertation. As such, I kindly request your permission to reproduce these items in my dissertation and related journal article (if applicable). Your reply to this email will be included in my dissertation's appendix as validation that I acted in accordance with copyright law. However, if I do not receive a response to this email by March 23, then I will take this to understand that you have given me passive consent to reproduce these rating scale items.

Thank you for considering my request. I'd be happy to answer any additional questions you might have regarding the reproduction of these items.

Best regards,

Danielle

Danielle Desfosses, PhD Candidate
Family and Human Development
T. Denny Sanford School of Social and Family Dynamics Arizona State University

Copyright Clearance Center Marketplace™

222 Rosewood Drive, Danvers, MA 01923 USA

INVOICE Page 1 of 2

Invoice No:	TRX20013691
Invoice Date:	10-Mar-2021
Account No:	3001834069
CCC Federal ID:	13-2922432
Terms:	Due upon receipt
Total Due:	10.76 CAD

Bill To:
MS. DANIELLE DESFOSSES
[redacted]

Ordered By:
DANIELLE DESFOSSES
[redacted]

Order Number:	1103276
Order Date:	10-Mar-2021

Republication Permissions

Project Name:	Not Available
Title:	Cascade Model of Executive Functioning, Prosocial Skills, and Academic Achievement
Authors:	Dr. Sabina Low

License ID	Description		Amount	
1103276-1	Type of use:	Republish in a thesis/dissertation	Price: CAD	10.76
	Publication:	Barkley Deficits in Executive Functioning Scale--Children and Ad...		
	ISBN-13:	9781462503940		

Payment should be made to **Copyright Clearance Center.**
Please see last page for payment instructions.

Subtotal:	CAD	10.76
Total Discount:	CAD	(0.00)
Total Due:	**CAD**	**10.76**

If you have any questions about this invoice, please email us at support@copyright.com or call us at Toll Free US +1-855-239-3415 / International +1-978-646-2600

APPENDIX E

INSTITUTIONAL REVIEW BOARD PERMISSION LETTER

April 9, 2021

To Whom It May Concern:

The Arizona State University Institutional Review Board (IRB) has been requested to determine if IRB approval was required for the use of a de-identified data set resulting from research performed under IRB approval “Second Step Efficacy Study” IRB 1208008096 directed by Sabina Low of The Sanford School. The study is now permanently closed and the data has been de-identified for future secondary use only.

Based on Dr. Low’s assertion that the data provided to Danielle A. L. Desfosses for her dissertation titled “Cascade Model of Executive Functioning, Prosocial Skills and Academic Achievement” was de-identified, the IRB has determined that IRB approval is not required.

The decision is based on the fact that secondary use of permanently de-identified data does not constitute Human Subjects Research under 46CFR46 because there was no interaction with any individual and no identifiable private information was used.

If you have questions or need more information let us know.

Sincerely,

Debra Murphy
Director, Research Operations
IRB Institutional Official
Knowledge Enterprise
660 S Mill Avenue Tempe AZ 85287-6111
web: research.asu.edu