Published in final edited form as:

Child Dev. 2019 September ; 90(5): 1614–1631. doi:10.1111/cdev.13091.

Developmental Transactions Between Self-Regulation and Academic Achievement Among Low-Income African American and Latino Children

Britain Mills, Nazly Dyer, Daniel Pacheco, Dawn Brinkley, Margaret T. Owen

University of Texas at Dallas

Margaret O. Caughy

University of Georgia

Abstract

This study examined the development of emerging self-regulation (SR) skills across the preschool years and relations to academic achievement in kindergarten and first grade. SR skills of 403 low-income African American and Latino children were measured at 2½, 3½, and 5 years (kindergarten). Reading and math skills were measured at 5 and 6 years (first grade) using the Woodcock–Johnson. Transactional relations between SR skills and achievement outcomes were estimated with latent difference score models. Increases in set shifting predicted prospective increases in reading, but not math scores. Increases in simple response inhibition predicted prospective increases in math, but not reading scores. Application of these findings to early intervention programming and needed supports for school readiness and achievement are discussed.

There is a growing body of literature linking selfregulation (SR) skills to school readiness and academic achievement in children (Blair & Razza, 2007; McClelland et al., 2007; Sektnan, McClelland, Acock, & Morrison, 2010). Defined as “a broad construct representing the cognitive, motivational-affective, social, and physiological processes that modulate attention, emotion, and behavior to a given situation/stimulus, for the purpose of pursuing a goal” (Bassett, Denham, Wyatt, & Warren-Khot, 2012, p. 597), SR skills support children’s ability to successfully navigate the demands of social and academic settings. Skills essential to success in the classroom such as not speaking out of turn, following directions, and managing frustration and other emotions all involve SR (McClelland, Geldhof, Cameron, & Wanless, 2015). Moreover, evidence that SR benefits from instruction and learning activities and predicts subsequent academic achievement has led to the development of multiple interventions aimed at improving SR and academic preparedness in early childhood (Bierman, Nix, Greenberg, Blair, & Domitrovich, 2008; Diamond & Lee, 2011; Raver et al., 2011).

Correspondence concerning this article should be addressed to Margaret Caughy, Department of Human Development and Family Science, University of Georgia, 108 Family Science Center II (House D), 405 Sanford Drive, Athens, GA 30602. margaret.caughy@uga.edu.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher’s website:

The Dimensionality of SR During Early Childhood

Despite growing recognition that SR plays an important role in children's academic, social, and emotional development, current understanding of how SR develops during early childhood and how this development is related to early academic achievement remains hindered by conceptual and methodological hurdles. The term "SR" is often used as an overarching term that encompasses both temperamental characteristics of the child as well as cognitive skills typically referred to as executive functions (EFs; Liew, 2012). At least three cognitive processes or EFs seem to be fundamental and include working memory (actively maintaining or manipulating information in memory), inhibition (suppressing a prepotent response), and cognitive flexibility or mental set shifting (SS), the ability to shift attention to a different aspect of the stimuli and respond accordingly (Blair & Ursache, 2011; Garon, Bryson, & Smith, 2008).

There also seems to be a general consensus that the dimensional structure of EF tasks exhibits both "unity and diversity" (Garon et al., 2008; Miyake et al., 2000). On one hand, the dimensionality of executive tasks is *diverse* in the sense that conceptually distinct constructs reflecting different neurocognitive mechanisms tend to cluster in accordance with the distinction between working memory, inhibition, and SS in factor analytic studies. On the other hand, executive tasks show *unity* in the sense that measures of specific EFs correlate, both with one another and with measures of a purportedly more general EF such as a central attentional or memory process (Posner & Rothbart, 1992).

Current understanding of the dimensionality of SR/EF comes largely from studies of older children and adults (see Miyake et al., 2000), whereas the dimensionality of SR/EF during early childhood remains an unsettled question. Willoughby, Pek, and Blair (2013) suggest that the consensus in the field is that EF in early childhood is best represented by a single dimension or factor that becomes multidimensional in later childhood. Although a single factor during early childhood was supported by the work of Wiebe, Espy, and Charak (2008), Wiebe et al. (2011), and Miller, Giesbrecht, Müller, McInerney, and Kerns (2012) identified a two-factor structure for EFs in early childhood (inhibition and working memory). A previously published report on the sample for this study at ages 2½ and 3½ supported a four-factor solution (Caughy, Owen, Mills, & Hurst, 2013). As noted by Miller et al. (2012), differences in factor structure can result from subtle differences in choices of tasks to assess SR/EF components. In addition, relying on a broad age range of children, such as in the studies reported by Wiebe et al. (2008) and Miller et al. (2012), may obscure developmental differences in the factor structure of SR/EF during early childhood.

In their systematic review of the literature on EF development during the preschool period, Garon et al. (2008) propose that executive processes develop in a hierarchical fashion, with precursor skills such as attention control developing first. This sets the stage for the subsequent emergence of working memory, followed by response inhibition, and ultimately SS. Garon et al. also review factor analytic and neuroimaging support for a distinction between simple and complex forms of response inhibition. Whereas the former involves inhibition of a single response, complex response inhibition tasks require participants to additionally activate another, conflicting response at an appropriate time. Complex inhibition

tasks are thus more demanding in that they require maintenance over a delay of an arbitrary rule specifying when a prepotent response must be suppressed and an incompatible response must be activated.

Questions about the differential validity of different SR/EF measures have been examined in a number of studies where measures of different SR/EF components are related to school outcomes. In this work, inhibitory control is consistently found to predict achievement (particularly mathematics, and sometimes reading; Bull, Espy, & Wiebe, 2008; Espy et al., 2004; Lan, Legare, Ponitz, Li, & Morrison, 2011). Findings for attention control (similar to SS) are less consistent, with some finding it to be a significant predictor of reading achievement (Lan et al., 2011), whereas others found no relation with academic achievement (Bull et al., 2008; Espy et al., 2004). Espy et al. (2004) argue that inhibitory control is a better predictor of young children's Woodcock-Johnson (WJ) Applied Problems performance, a commonly used measure of mathematics ability, because the earlier items on the test are simple problems that depend on knowledge of small-quantity numbers.

Longitudinal Studies of SR During Early Childhood

As noted by Garon et al. (2008) current understanding of the development of SR/EF components is based on cross-sectional analyses of children at different ages. The few studies that have studied EF longitudinally have relied on a single indicator of EF and/or restricted the range of the longitudinal period to < 3 years (Fuhs, Nesbitt, Farran, & Dong, 2014; Hughes, Ensor, Wilson, & Graham, 2010; Schmitt, Geldof, Purpura, Duncan, & McClelland, 2017; Welsh, Nix, Blair, Bierman, & Nelson, 2010; Wiebe, Sheffield, & Espy, 2012). It is also unclear how the different components of SR are structured together in the early years of development. As noted in a review by Best and Miller (2010), how the onset or course of particular executive processes affects others remains largely a mystery. This knowledge gap extends to theorized correlates such as academic achievement. Currently, little is known about how the onset and development of individual executive processes prospectively affect different aspects of achievement. Longitudinal studies of SR's influence on various outcomes have been conducted, but they have typically modeled the outcome longitudinally and treated SR as a time-invariant covariate (Blair, Ursache, Greenberg, & Vernon-Feagans, 2015; Li-Grining, Votruba-Drzal, Maldonado-Carreño, & Haas, 2010; Matthews, Kizzie, Rowley, & Cortina, 2010; McClelland, Acock, & Morrison, 2006). This approach precludes the possibility of detecting dynamic associations between SR and the outcome over time in the form of correlated concurrent changes or prospective transactional effects of one variable on another.

A few longitudinal studies of SR /EF development have included repeated measures of both SR/EF development and academic achievement outcomes including Fuhs et al. (2014), Schmitt et al. (2017), and Welsh et al. (2010). Although similar in purpose to this study, they differed in some important ways. For example, each used a composite measure of EF rather than measures of distinct SR/EF components. This precluded examining dynamic relations between different components of SR/EF and achievement outcomes. In addition, these studies focused on children who were older, starting when children were in prekindergarten (around age 4½ years), unlike the children in the present investigation who were 2½ years

old at study inception. Modeling SR/EF longitudinally starting at such a young age—when SR/EF skills are emerging and developing rapidly—provides a unique opportunity to examine both the codevelopment and dynamic interactions between these skills and achievement, as children transition from a critical formative period to formal schooling.

SR Development Among Ethnically Diverse Children During Early Childhood

Another limitation of previous work is that the sociocultural composition of extant samples is largely non-Latino white and middle class (e.g., 80%–95% among studies reviewed in Garon et al., 2008 that reported the ethnic composition of their sample). One notable exception is the The Family Life Project (see Rhoades, Greenberg, Lanza, & Blair, 2011; Willoughby, Wirth, & Blair, 2012), a recent study that includes a substantial number of low-income African Americans. However, measures of SR constructs were not administered prior to age 3 in this study. To understand how the capacity for SR emerges, it is clear that more data on diverse samples are needed. This study focuses on an underrepresented population in the SR literature, specifically low-income African American and Latino children, the latter being primarily of Mexican-origin and Spanish-speaking when the study began. Although there is no theoretical reason to expect that the development of SR/EF would proceed differently in the current sample, increasing the diversity of samples studied in the literature on EF development is important for several reasons. For example, previous analyses of the present data indicate that the milestones of achieving proficiency in the component skills of SR/EF are achieved significantly later as compared with what has been reported previously (Caughy et al., 2013). Furthermore, because low-income ethnic minority children are at heightened risk for early academic failure relative to the samples of children dominating the preschool EF literature, studying these developmental processes within this population is critical for the development of interventions to support early academic achievement.

This study directly addresses many of the limitations of the available research on the development of SR/EF during early childhood and its relation to early academic achievement. Performance data on SR/EF tasks and academic achievement are examined in a sample of children followed longitudinally starting at age 2½. Additionally, it focuses specifically on African American and Latino children from low-income families, a high-risk and understudied population of children in the EF literature. African American and Latino children from low-income families experience higher rates of early academic failure, and poor SR development has been suggested as one of the factors that may at least partially mediate the relations between poverty-related stressors such as household instability and parenting stress and poor academic achievement (Raver, 2004; Ursache, Blair, & Raver, 2012). However, studies of SR development in this population are exceedingly rare, particularly among Latino children and particularly starting at such a young age.

A Latent Difference Score Framework for Studying SR Development

Longitudinal transactions between multiple latent SR/EF variables and achievement measures were modeled using a latent difference score (LDS) framework (Ferrer &

McArdle, 2004; McArdle, 2009). This approach reflected several considerations. For example, given general expectations that SR will exhibit developmental increases, models of within-person change are warranted. As will be discussed here, this is doubly important given that in previous work, *change* in the capacity for SR has often not been distinguished from the absolute capacity for SR. Latent growth modeling represented an alternative approach that was not used due to limits on available assessments (many of the SR variables were assessed at two adjacent time points) and the analytic goals. Specifically, the goals were not to describe the functional form of developmental trajectories of SR variables. Rather, the purpose was to model dynamic transactions between SR components and their prospective impact on theoretically relevant outcomes. LDS models are ideally suited for these questions and this methodological context.

As shown in Figure 1, the LDS model is simply a reparameterized version of the familiar crosslagged regression approach to longitudinal analysis. By imposing a few constraints, however, a crosslagged model of latent variables over time can be reframed as a model of latent *change* in those variables over time. For example, latent change in a longitudinal variable Y (Y) over two time points can be obtained by constraining (a) the means/intercepts of Y_1 and Y_2 to equality, (b) the regression of Y_2 on Y_1 to 1, (c) the residual variance of Y_2 to 0, and (d) the regression of Y_2 on Y to 1. The mean/intercept and residual variance of Y is freely estimated. This reparameterization is analogous to the common practice of directly computing a change score (i.e., $Y_2 - Y_1$). However, because change is computed indirectly through model constraints in the LDS model, strengths of the latent variable framework are retained (e.g., more reliably estimated structural parameters, and in turn, greater power to detect effects). The LDS approach simultaneously avoids common complications of stepping outside of the latent variable framework, such as unreliability in manifest change scores, or the often problematic use of factor scores to compute such change scores (Cronbach & Furby, 1970; Grice, 2001).

Parameters of an LDS model directly quantify various within-person effects of possible theoretical interest and are thus well-suited for modeling transactions between rapidly changing variables. The latent change score's intercept corresponds directly to the average amount of within-person change in the latent variable between two assessment points, and effects on the latent change variable represent unique effects on this within-person change. In contrast, in the cross-lagged regression approach, change is modeled only indirectly, and its parameters quantify various *between-person* effects. For example, autoregressive parameters describe how much the rank ordering of a variable's values reshuffle across time (an overview and comparison of several major alternative approaches to longitudinal data analysis with structural equation models can be found in Selig & Preacher, 2009).

The LDS model is similar to the more familiar latent growth model (LGM) in that both can be used to model change over time. A key difference between these approaches is whether time is explicitly parameterized. In the LGM, a latent slope factor drives linear change over time. In the LDS model (as well as autoregressive/cross-lagged models), time is accommodated implicitly by incorporating it directly into the definitions of the variables. The stable component of the trajectory is transmitted through fixed autoregressive paths between adjacent time points, and the trajectory's dynamic components are freely estimated

as deviations from their value at the preceding time point. Because no constraints are imposed on these deviations, they freely accumulate at each time point and may take any functional form, linear or nonlinear. As discussed by McArdle (2009) the LGM is a special case of the LDS model and is particularly useful when one wishes to impose functional form on a variable's change over time through the estimation of temporal trajectory parameters such as intercepts and slopes. A focus on the functional form of the change makes sense with three or more data points. However, when some measures are only available for two time points but one still wants to model change between those two time points, such as in this study, or when describing functional form is not an overarching goal, then the LDS model provides greater analytic flexibility than does a LGM approach.

In addition to elucidating the emergence of SR skills during early childhood, another purpose of this article was to examine the prospective relations between these emerging SR skills and academic achievement after the transition to formal schooling. As noted previously, few studies have examined the longitudinal changes in SR skills, and even fewer have examined the relation of change in SR skills to academic achievement. This study adds to previous work by including direct, repeated assessments of multiple SR skills during early childhood, a period when these skills emerge and exhibit rapid growth. This provides a unique opportunity to examine how individual differences in the emergence and subsequent developmental course of these early childhood skills predict academic functioning. This question is fundamentally different—theoretically and methodologically—from those addressed in previous research, which have focused almost exclusively on how SR skills at a single point in time (such as during preschool or kindergarten) predict later academic achievement or growth in achievement. The nature of the relations between emergent SR skills, dynamic changes in those skills, and subsequent academic outcomes in this sample of at-risk children has important implications for the development and timing of early interventions to support academic success. For example, if certain components of EF are more predictive of particular academic outcomes than others, or if onset and change are differentially predictive of subsequent achievement, knowledge of these facts can directly inform the content of interventions and the efficiency with which they are delivered.

Method

Participants

A total of 407 children and their families were recruited in 2010 as part of a large longitudinal study examining SR development and school readiness among low-income African American and Latino preschoolers living in a large metropolitan area in the southwest United States. Families were recruited from the community through the dissemination of project information to organizations such as WIC clinics and Head Start programs, recreational centers, churches and by word of mouth. To be eligible for the study, at least one parent had to self-identify as either African American or Latino with family income that was below 200% of the federal poverty level. At the time of enrollment, recruited families had one child who was between 29 and 31 months old who was not hospitalized for more than 7 days after birth. Additionally, the family needed to intend on

remaining in the area for at least 1 year. Four waves of data collection have been completed: at age 2½, age 3½, kindergarten, and first grade.

Of the initial 407 families enrolled in the study, four were excluded from the analysis because of subsequent diagnoses of autism or other significant developmental disability. The remaining 403 children (54% boys, 46% girls) were retained for the current sample, of which 45% were African American and 55% were Latino. This study examines SR data from Waves 1–3 in relation to child outcomes during Waves 3 and 4. Of these 403 children, 32 (7.9%) were lost to follow-up after Wave 1, 39 (9.7%) completed two waves, 33 (8.2%) completed three waves, and 299 (74.2%) completed all four waves of data collection. In addition, 333 (82.6%) completed data collection during Wave 3 and/or Wave 4, when academic achievement was assessed. Likelihood of follow-up did not differ by child gender or child SR skills at Wave 1, but there were differences by child ethnicity and household income. African American children were more likely to be lost to follow-up after one visit compared to Latino children, 12.1% versus 4.5%, $\chi^2(1) = 7.81, p < .01$. In addition, families lost to follow-up had lower average income-to-needs ratios compared to those who were not. The average income-to-needs ratio for families who completed one or two waves of data collection was .62 ($SD = .53$) compared to .86 ($SD = .55$) for those who completed three or more, $t(396) = 3.64, p < .001$.

Sample characteristics are reported in Table 1. There were several key differences between African American and Latino families. The primary care-giver was more likely to be the biological mother among Latino families compared with African American families. Average family income was higher for Latino families, and African American caregivers were more likely to have a high school education or higher compared to Latino caregivers. Latino fathers or father figures were more likely to be living in the household compared to African American fathers or father figures.

Data Collection Procedures

Home visits lasted approximately 1.5 hr and were conducted by a team of two home visitors. One home visitor interviewed the primary caregiver, whereas the second home visitor conducted the child assessments. Although all child assessments were video-recorded, some tasks were coded live in the home and others were coded later in the lab, as will be described next. Home visitors for all Spanish-speaking families were bilingual. All measures were administered to parents and children in their preferred language.

Measures

Child SR and EF—Key features of children's SR skills were measured including inhibition, SS, and working memory. Inhibition was measured by two types of tasks, simple response inhibition (inhibition tasks requiring minimal load on working memory and the inhibition of a single response), and complex response inhibition (tasks which put greater demands on working memory in that they involve multiple conflicting response options; Garon et al., 2008; Petersen, Hoyniak, McQuillan, Bates, & Staples, 2016).

Simple response inhibition.: Simple response inhibition was assessed using four tasks: Snack Delay (Waves 1 and 2), Wrapped Gift/Wait for Bow (Waves 1–3), Forbidden Toy (Wave 1 only), and Delay of Gratification (Wave 3 only). During the *Snack Delay* task (Kochanska, Murray, & Harlan, 2000), children were instructed to wait until the home visitor rang a bell before eating a small chocolate candy placed in front of them. The task consisted of four trials (10 s, 20 s, 30 s, 15 s) where latency (measured in seconds) from the start of the trial until the child touched the candy was coded from video recordings. Interrater reliability based on intraclass correlation (ICC) for the snack delay task coding was .99 based on double coding 15% of cases.

During the first phase of the *Wrapped Gift/Wait for Bow* task (Kochanska et al., 2000), the child was instructed not to peek while a gift was wrapped directly behind them for 1 min. During the second phase of the task, the child was instructed not to touch the gift that was placed in front of the child, whereas the child assessor left the room for 1½ min to get a bow. Latency to peek was coded during the wrap phase, and latency to touch the wrapped gift was coded during the wait for bow phase. Interrater reliability (ICC) for the wrapped gift task ranged from .87 to .99 across waves based on 21% double coding of cases. During Wave 3, only latency to peek was coded.

The *Forbidden Toy* task, adapted from procedures of the NICHD Study of Early Child Care and Youth Development (NICHD ECCRN, 1998) was only administered during Wave 1. The home visitor and child played with the attractive toy (a small car that moved on its own after shaking it) for a full minute. After the 60-s play period, the home visitor told the child not to touch or play with the toy until the visitor returned to the room. The wait trial lasted 150 s. Latency to touch was scored from videotape. Interrater reliability for the task was high (ICC = .95) based on 16% double-coded cases.

During Wave 3, the traditional *Delay of Gratification* task based on the classic work of Mischel and colleagues was used (Mischel, Shoda, & Rodriguez, 1989). The child was shown two piles of snacks (either M&Ms or gold fish crackers), one smaller, and the other larger. The child was told if s/he waited for the entire 6 min, s/he would get the larger pile. The child was told s/he could ring the bell sooner, at any time during the waiting period, to get the smaller pile. Latency to ring the bell was coded. Interrater reliability was high (ICC = .92) based on 20% double-coded cases.

Complex response inhibition.: Complex response inhibition was assessed using two tasks, Fruit Stroop (Wave 1) and Head-Toes-Knees-Shoulders (HTKS; Waves 1–3). In the *Fruit Stroop* task (Kochanska et al., 2000) children were presented with pictures of large fruits with smaller, different fruits embedded in them. In order to successfully complete the task, the child needed to inhibit his or her dominant response to point to the large fruit and point to the small fruit instead. Children were scored on their ability to name each fruit (two trials), correctly point to the small fruit in practice trials with reinstruction for incorrect responses (two trials), and correctly point to the small fruit during test trials (four trials). Each trial was coded on a 3-point scale (0 = *incorrect*, 1 = *self-correct*, 2 = *correct*).

The HTKS task was developed by McClelland and colleagues (Ponitz et al., 2008). During Wave 2, only the first portion of the task (Head & Toes) was administered due to prominent floor effects. In each trial, the child was asked to touch his or her head when the child assessor said to touch and did touch her toes and vice versa. The task consisted of six practice trials and 10 test trials, each coded on a 3-point scale: *incorrect* (0), *self-correct* (1), and *correct* (2). During Wave 3, all three phases of the HTKS shoulder were administered. Phase 2 (Knees and Shoulders) included an additional five practice trials and 10 test trials. In the final phase of the task, all four body parts are included, but the corresponding body parts are changed so that head was paired with knees and toes were paired with shoulders. Phase included six practice trials and 10 test trials for a total of 30 test trials across all three phases. Interrater reliability (ICC) for Wave 2 based on 16% of cases double-coded was .89; for Wave 3, interrater reliability was .90 based on 21% of double-coded cases.

Set shifting.: SS was assessed during Waves 2 and 3 using the Dimensional Change Card Sort (DCCS) task (Diamond, Carlson, & Beck, 2005). The task involves multiple sets of cards, each containing cards with fish and with cars in red or blue and a pair of sorting boxes, one labeled with a red fish and the other with a blue car. During the warm up and priming trials, the child was instructed to sort the cards into the boxes based on one criterion (either color or shape). For the test trials, the child was asked to switch rules and sort the cards by the other dimension, instructing the child that the game had now changed from a “shape game” to a “color game,” or from a “color game” to a “shape game.” Of the six postswitch trials, four were conflict trials that required the child to consider both the shape and color of the card’s pictured object in order to sort the card by the correct dimension. This was done in order to ensure that scores for correct SS were not due to chance. The dimension used for priming trials versus test trials (color or shape) was counterbalanced across children. Performance on the conflict trials was scored as either incorrect (0) or correct (1). Interrater reliability (ICC) based on 24% double coded was .92. In Wave 3, an extension of the DCCS was administered to include a border phase. In the border phase, children were asked to sort the card by color if the card has a black border and to sort by shape if the card does not have a black border. There were 12 test trials during the border phase. Interrater reliability (ICC) based on 21% double-coded was .90 in Wave 3.

Working memory.: A memory span task called *Memory Chocolates* was created for the study using a toy called “Smart Snacks Hide “n” Peek Chocolates” by Learning Resources, a heart-shaped plastic box with removable lid and spaces for twelve differently shaped plastic “chocolates.” Each chocolate had a removable top and various animal stickers were placed with a clear adhesive coating on the plastic “chocolate” underneath the top. The task was administered during Wave 2 and consisted of six naming trials, during which the covers were removed, and the child was asked to name each of the animals, after which the covers were replaced. Three practice trials in which the child was asked to find a single animal followed. The child had to pass at least two practice trials to continue to the test trials. Three test trials assessed a memory span of two items, and a second set of test trials assessed a memory span of three items. The task was scored live. During Wave 3, the Operation Span (OS) task (Willoughby, Blair, Wirth, Greenberg, & The Family Life Project Investigators, 2010) was administered. In this 19-item working memory task, children had to remember

types of animals displayed on a previous page in the correct order. The task required an increasing number of items children had to remember at a given time, from 2 to 4 animals. This task was scored live during the assessment.

Academic achievement outcomes.: To measure academic outcomes, the Applied Problems, Letter-Word Identification, and Word Attack subtests of the standardized WJ achievement battery (Woodcock, 1990) were administered during Waves 3 and 4. For children who were Spanish-dominant, the Spanish version of the test, the Bateria Woodcock–Muñoz (Woodcock & Munoz-Sandoval, 1996), was administered.

Specific guidelines were followed for determining the child's language preference and the form of the WJ achievement battery (WJ reading [WJR]) used in the home visits. This was done through the Receptive One Word Picture Vocabulary Test (ROWPVT; Gardner, 1985), which was also administered but is not included in the current analyses. The ROWPVT includes a bilingual version and extensive questions about the child's language preferences, providing a clear idea of the child's language proficiency in English and Spanish before administering the WJR. Only a single language could be used in each subtest's administration, but for the Applied Problems subtest, the child could answer questions using either language. For a correct score in the Letter-Word Identification and Word Attack subtests, the child needed to answer in the same language as the language used in the test's administration.

W scores were used for these analyses. For all models, the WJR scores were centered at their normed mean of 500 and rescaled by dividing the resulting value by 20. This rescaling reduces the estimated variances for WJR scores to values more consistent with scaling of other variables in the model, in order to avoid common convergence problems (Muthén & Muthén, 1998–2006). A one point change in the rescaled metric thus corresponds to a 20 point change in the original WJR metric.

Performance on the Applied Problems subtest was used as a measure of achievement in mathematics, and the Letter-Word Identification and Word Attack subtests scores were averaged for a composite measure of reading achievement.

Covariates.: Covariates used in the analyses reported included child ethnicity and family income-to-needs ratio. For the latter, the ratio of reported household income to family size at each wave was averaged.

Translation of Study Materials

All measures that were only available in English were forward translated into Spanish and then back-translated into English by two Spanish-speaking individuals. Differences were resolved by consensus.

Analyses

To inform the initial specification of a longitudinal model of transactions between SR latent variables, several preliminary analyses were conducted. First, confirmatory factor models were estimated for SR measures administered at each wave of assessment based on a

previously reported analysis of the dimensionality of these measures at Waves 1 and 2 (Caughy et al., 2013). As in those previous analyses, all SR latent variables were modeled using trial level data, and all trial level data were dichotomized. For SR tasks in which trial data were latencies, these latencies were dichotomized as did not wait (0) versus waited (1) given that the latencies exhibited distinct bimodal distributions where the large majority of responses fell into the two most extreme response categories. The categorical scale of the indicators was accommodated by modeling their factor loadings as probit coefficients through the use of Mplus mean and variance-adjusted weighted least squares estimator (Muthén & Muthén, 1998–2006).

In previously reported analyses of SR data from Waves 1 and 2 of this study, systematic confirmatory factor analyses comparing single-factor, two-factor, three-factor, and four-factor representations of the data provided support for a four factor model (Caughy et al., 2013). Results of wave-specific analyses of all four waves of SR data were largely consistent with a four factor model of SR including latent variables to represent simple response inhibition, complex response inhibition, set shifting, and working memory. Subsequent longitudinal variations on these models employed the same approach, where SR measures were partitioned into four substantive sources of variation. The amount of variance explained in the SR measures by their respective factors was generally high. Indicator R^2 values for the four substantive SR factors ranged from .63–.91 (SS), .08–.90 (simple response inhibition), .49–.94 (complex response inhibition), and .06–.84 (working memory).

Next, SR measures administered across multiple time points were subjected to a series of nested model comparisons to test for longitudinal measurement invariance, using time of assessment as the grouping variable in a multiple-group framework (Steenkamp & Baumgartner, 1998; Vandenberg & Lance, 2000; Widaman, Ferrer, & Conger, 2010). These analyses were conducted separately by SR measure (e.g., simple response inhibition items only), and their purpose was to test whether measurement parameters (e.g., loadings, thresholds) differed by time of assessment. When measurement differences are present and are not addressed in some way such as by removing the offending items, or by directly incorporating measurement differences into the model, they can bias inferences about the construct of interest. In longitudinal contexts, the concern is with the possibility that change in the construct is confounded with change in the construct's measurement. In the present analyses, strong measurement invariance as reflected by no differences in loadings or thresholds across time held for all longitudinal measures (for additional details, see Table S1). Analyses establishing measurement invariance by gender and ethnicity were reported in Caughy et al. (2013).

In an LDS model, proportional change represents the extent to which change depends on the starting/baseline value, or the effect of $X(t-1)$ on $X(t)$ and is analogous to the time by baseline score interaction effect commonly estimated in models that explicitly parameterize effects of time (see Figure 1b). Negative proportional change is common and simply indicates that more extreme scores at an earlier time point tend to be followed by larger changes *in the opposite direction* at the following time point. This may reflect some substantive aspect of the developmental process being modeled, a regression to the mean artifact, or some combination of these two explanations. Of course, regression to the mean is

present any time repeated measurements are made on scores with random error, and disentangling it from substantive developmental processes presents well-known difficulties (Lord, 1967). Whatever its source(s) however, when present it cannot be ignored. Significant, nonzero estimates of proportional change reflect a real dependency in the data that can distort other estimates of interest if it is not adjusted for.

Our general modeling strategy was to first estimate a model with saturated structural effects over time. That is, all possible lagged effects from prior time points were specified as in Figure 1b. In a subsequent stage of model trimming, nonsignificant effects were dropped, and models were refit. Trimmed models differed in no substantive way from the initial model with saturated effects, and consequently, the initial models specified prior to any exploratory model respecifications are reported. Figure 2 provides an overview of the latent variables in the model, the corresponding measures they were based on in parentheses, the waves at which they were administered, and exactly which variables were parameterized as LDS in the analyses. Note in particular that latent variables for working memory and the Fruit Stroop task were not parameterized as latent change scores. As described in the introduction, modifications to the measurement instruments were necessary over the course of the study. Fruit Stroop and Memory Chocolates were two measures that were not readministered in a subsequent wave. Although an alternative measure of working memory, Operation Span, was added at Wave 3, this was an entirely different measure with no items in common with Memory Chocolates. Consequently, it was not sensible to compute any kind of change score either directly or indirectly for working memory. However, because Memory Chocolates and OS reflect the same theoretical construct, these measures are grouped together in Figure 2 in a manner analogous to other latent variables in the figure. Fruit Stroop is modeled as a distinct latent variable, as its items consistently clustered together in preliminary factor models.

Models were estimated separately for reading and math achievement scores on the WJ. All reported models additionally adjusted for effects of the child's ethnicity and family income-to-needs ratio on the achievement outcome at Wave 3 and the latent change in achievement at Wave 4.

Results

Results from two LDS models are shown in Table 2. The two models differed only in the achievement measure (reading vs. math) modeled at Wave 4. Structural estimates not involving these outcomes such as SR means and paths linking SR variables are shown from the reading model and differed in no substantive way from the math model.

LDS include those variables with γ in the name. As can be seen in the first column on the far left, the means/intercepts for all LDS were consistently positive and highly significant. For example, between Waves 1 and 2, scores on the simple response inhibition measure increased by roughly 1.7 *SDs*. Likewise, scores on the WJR reading and math measures improved between Waves 3 and 4 by just over 4 points and 3 points, respectively. Both of these results make sense theoretically because the capacity for SR should increase for

children over this age period as should academic achievement among children receiving academic instruction.

For all latent difference variables, proportional change was significant and negative. As described earlier, this indicates that higher proportional change scores at earlier time points tended to be followed by higher proportional change scores in the opposite direction at the next assessment. This is a common finding, and all structural estimates of interest discussed next control for this dependency on starting values.

Prospective Effects of Emerging SR Skills

One of the goals of this article was to understand developmental transactions between emerging SR skills. Transactional relations between SR measures are shown under the heading of “SR (Wave 3)” in the right half of Table 2. As described in the Table 2 footnote, prospective effects on Wave 2 SR measures were nonsignificant and are not shown in the table. Also not shown in the table are correlations between the exogenous Fruit Stroop and simple response inhibition variables ($r = .36, p < .001$) and correlations between endogenous residual variances variables at the same assessment point. Significant correlations among these latter estimates included Wave 2 CI with SS ($r = .25, p < .05$), CI with SS ($r = .34, p < .05$), CI with OS ($r = .44, p < .01$). Wave 3 SR variables were also correlated with Wave 3 achievement outcomes: CI: WJR: $r = .60, p < .01$; WJ math (WJM): $r = .36, p < .001$; SS: $r = .48, p < .05$; WJM: $r = .24, p < .05$; and OS: $r = .70, p < .001$; WJM: $r = .27, p < .01$. All correlations are provided in Table S2.

Significant prospective estimates shown in Table 2 included a positive effect of Fruit Stroop at Wave 1 on later changes in complex response inhibition, positive effects of simple response inhibition on subsequent changes in SS and in working memory (as measured by the OS task), and a positive effect of changes in simple response inhibition on later changes in SS. There was also a positive effect of Wave 2 working memory (as measured by the Memory Chocolates task) on later changes in SS. An asymmetry within these transactional relations to note is that in general, simple response inhibition and working memory measures had positive prospective influences on the later SR measures of SS and complex response inhibition, but the converse was not the case. That is, SS and complex response inhibition did not show any prospective effects on one another or on working memory.

Prospective Effects of Emerging SR Skills on Academic Achievement

The second purpose of this article was to examine the prospective effects of emerging SR skills on academic outcomes in early elementary school. The right half of Table 2 contains effects of SR measures on later achievement outcomes. There were two sets of significant parameters, and again there is an interesting asymmetry in that the pattern of effects varies across the two achievement outcomes. First, increases in SS predicted increases in reading achievement, as indicated by the significant, positive estimate for the $SS \rightarrow WJR$ reading path. Second, children with higher general levels of simple response inhibition at Wave 1 tended to score higher in general on math achievement. Over and above this effect, increases in simple response inhibition between Waves 1 and 2 also predicted increases in math achievement between Waves 3 and 4.

Discussion

During the preschool period, self-regulatory skills develop along a steep and accelerating arc. These abilities tend to show remarkable improvements over a relatively short time as children begin to rely increasingly on their own internal cues and less on external controls provided by caregivers (Kochanska, Coy, & Murray, 2001). Surprisingly, however, only a handful of studies have sought to estimate how different dimensions of SR are rapidly changing, how changes in these dimensions relate to one another, or how they relate to relevant achievement outcomes. In this study, an analytic approach tailor-made for these types of questions—LDS models—was used to model transactional relations involving four SR dimensions across early childhood and the transition to school and reading and math outcomes. Results showed a distinct asymmetry in transactions. SR skills thought to emerge earliest—working memory and simple response inhibition—displayed prospective effects on later emerging skills such as complex response inhibition and SS. However, these more advanced EFs of complex response inhibition and SS did not demonstrate prospective effects on working memory or simple response inhibition. These findings indicate that working memory and simple response inhibition represent foundational processes that support the subsequent emergence of more sophisticated functions like SS. They also echo findings of Friedman, Miyake, Robinson, and Hewitt (2011), who showed that individual differences in simple response inhibition during early childhood predicted individual differences in SS in adolescence. This asymmetrical development of SR/EF skills over the course of early childhood supports the hierarchical model proposed by Garon et al. (2008).

The present results also support a more nuanced view of the dimensionality of SR during early childhood in that distinct dimensions of SR related differently to math and reading achievement outcomes. Invariance testing across ages 2½, 3½, and kindergarten indicated four distinguishable components of SR—simple response inhibition, complex response inhibition, working memory, and set shifting—consistent with a previously published report (Caughy et al., 2013). The hierarchical relation between emerging SR skills, wherein early emerging skills prospectively affect later skills, lends further support to a view of distinguishable components of EF during the preschool years. The present findings regarding dimensionality may differ from previous results for many reasons. One is that the type as well as the breadth of measures used to assess SR/EF can affect the number of dimensions found (Miller et al., 2012). For example, measures with an affective component such as snack delay often load on a separate dimension such as in this study (Bassett et al., 2012; Denham, Warren-Khot, Bassett, Wyatt, & Perna, 2012). Age of the child during assessment is also likely to be an important factor, as the dimensional structure of SR/EF may change over the course of development. Children in this study were not only chronologically younger at enrollment than other studies reporting on the dimensionality of SR/EF in early childhood, they also performed at levels that were significantly below (about 8 months behind) reports of similarly aged samples (Caughy et al., 2013).

A second purpose of the present investigation was to examine the prospective effects of emerging SR skills on academic achievement outcomes in early elementary school. The results of the LDS analysis indicated that increases in SS were prospectively related to increases in reading achievement between kindergarten and first grade. Higher levels of

simple response inhibition at age 2½ were related to higher levels of mathematical ability, and increases in simple response inhibition between ages 2½ and 3½ also predicted increases in mathematical ability. This pattern of effects for math is perhaps somewhat puzzling given previous accounts of the kinds of processes engaged by mathematics. According to Blair et al. (2015; see also McClelland et al., 2014), a mathematics problem represents a “classic” EF task because it requires switching attention between relevant problem elements. Under this line of thought, SS might be expected to be a key predictor of mathematical ability. Why this effect did not emerge in the present analyses is unclear. One possibility is that SS is simply not engaged by the outcome measure at this age among these participants. For example, because early items on the Woodcock-Johnson rely on more basic knowledge of numeracy (Espy et al., 2004), the measure may be less likely to engage SS processes. To the extent the processing demands of the task are restricted to application of basic knowledge, simple inhibitory processes may be more relevant to successful performance.

Some researchers suggest that it is important to statistically control for processing speed or general IQ when examining the relation between executive cognitive skills and academic outcomes, although this practice has been questioned (e.g., Cepeda, Blackwell, & Munakata, 2013; see also Dennis et al., 2009). At issue are questions about how or whether such covariates are situated along the causal pathways that link neurocognitive processes to performance outcomes such as academic achievement. For example, if covariates like IQ or processing speed are causal consequences of self-regulatory processes such as inhibition or working memory, then controlling for those covariates in statistical models of performance outcomes would logically be expected to distort the very parameters of interest (namely, those that estimate the causal effect of a self-regulatory process on a performance outcome). With respect to the present findings, the issue was inconsequential: Supplemental analyses of the present data that included an adjustment for processing speed had no substantive impact on the reported results (results available upon request).

Regarding executive functioning and reading, Blair et al. (2015) suggest that the ability to identify smaller units within larger words would require both attention shifting as well as inhibitory control skills. However, Blair et al. (2015) contend that early literacy is based more on knowledge of letters, sounds and words and should not depend on executive processing skills. In this study, although SS skills were not associated with reading achievement in kindergarten, increases in SS between age 3½ years and kindergarten predicted increases in reading skills from kindergarten to first grade. It may be that the growth in reading skills between kindergarten and first grade represents the emergence of the higher order literacy skills of phonological awareness and decoding skills that Blair et al. (2015) hypothesized would be facilitated by a child’s set shifting ability.

Several other studies have examined the differential relation of specific components of SR with academic outcomes, but the findings across studies are not consistent enough to draw any firm conclusions. Comparisons across studies are complicated by whether multiple dimensions or a single dimension of EF was assessed, the variety of measures employed, and whether EF measures were combined in the analysis or considered separately. Among researchers who have examined relations of individual components of EF and later academic

achievement, some have found that measures of complex response inhibition are related to math performance but not reading (Brock, Rimm-Kaufman, Nathanson, & Grimm, 2009; Lan et al., 2011; Matthews, Ponitz, & Morrison, 2009; McClelland et al., 2014; Ponitz, McClelland, Matthews, & Morrison, 2009), whereas others have reported that complex response inhibition is related to both math and reading performance (Bull et al., 2008).

This study differed from other studies of child SR and achievement in several ways. For example, SR was modeled longitudinally, as opposed to examining the relation between SR assessed at one time point and later academic achievement. Longitudinal studies of SR's influence on academic outcomes have typically modeled the outcome longitudinally, but not SR (Blair et al., 2015; Li-Grining et al., 2010; Matthews et al., 2010; McClelland et al., 2006). Additionally, different components of SR/EF were measured beginning when children were 30 months old. The few studies that have modeled both SR as well as academic achievement longitudinally utilized a composite of SR/EF as opposed to modeling the components of SR/EF separately, and started at a significantly older age compared to this study (Fuhs et al., 2014; Schmitt et al., 2017; Welsh et al., 2010). Together, these aspects of the present methodology allowed the stable and dynamic contributions of different SR skills to achievement to be disentangled over the course of early childhood.

Examining emerging SR longitudinally during this period of rapid development presents a number of methodological challenges. As noted by Garon et al. (2008), although the EF skills of young children cannot be assessed using the same tasks used with older children/adults, simply modifying tasks used with older individuals is risky because it remains unclear whether the critical EF skill measured by the task is still being assessed with validity. Furthermore, because cognitive skills are developing rapidly during early childhood, EF tasks have short windows of utility, defined as the age range when the measure will have sufficient range and variability. At ages earlier than this range, the measure will suffer from large floor effects, and after it, ceiling effects are a problem. Petersen et al. (2016) conducted a meta-analysis of complex response inhibition tasks between the age of 18 months and 8 years and found that measures were useful for a developmental span of < 3 years. However, it is important to note that this limitation is not unique to this study: it is intrinsic to the study of any emergent construct. Future research should continue employing developmentally sensitive tasks while being mindful of task congruency over time, at least to the extent it is possible without inducing floor or ceiling effects. A second limitation was the relatively small number of assessment points. As with any longitudinal study, more assessment points would allow for more fine-grained analysis of the processes under examination. Finally, the focus on low-income African American and Latino children is both a strength and limitation of this study. While this population is understudied, the focus on specific ethnic groups from low-income families limits the generalizability of the findings to other groups.

The present focus on a relatively understudied population in the early childhood EF literature—low-income African American and Latino children—is notable in that no other study of SR development during early childhood has included a substantial number of Latino children. Although the Family Life Project includes a large number of low-income African American children, there is still a need for additional study of the development of

SR and its components with data on more diverse samples. The current findings indicate that early SR skills support later academic achievement among low-income African American and Latino children. Consequently, improving understanding of the development of SR in these populations is important for supporting early academic achievement in these groups.

The relevance of SR to later academic achievement, however, has been questioned in at least one study. Willoughby, Kupersmidt, and Voegler-Lee (2012) argued that the relation between early EF development and later academic achievement is spurious because the relation is confounded by time-invariant, unmeasured covariates. This was inferred from the disappearance of an EF-achievement association in a fixed effect regression model that—unlike other longitudinal analytic techniques (the authors argued)—controlled for time-invariant unmeasured covariates. However, the fixed-effect model is a special case of a more general regression model with specific, restrictive assumptions about random effects and time-varying influences, and is typically used when one has no intention of generalizing to other time points. Consequently, to the extent the fixed effect model controls for such effects, so does any more general regression-based approach, including the LDS approach used in the current investigation. As such, the present finding that early SR predicts later academic achievement is also adjusted for the potential confounders highlighted by Willoughby, Kupersmidt, et al. (2012), suggesting the association is not spurious.

This study is the first to examine longitudinal relations between emerging EFs and academic skills in a low-income sample of African-American and Latino children. Minority children have been shown to be at risk for the development of deficient EF skills (Raver, 2004) and academic failure (U.S. Department of Education, 2004). Additionally, ethnic minority children disproportionately come from low-income households, which has been robustly associated with a variety of suboptimal outcomes. Consequently, it is important to examine the manner in which regulatory competencies develop and relate with academic achievement in populations that have been historically underserved and understudied. Results from this study are a first step in understanding transactions between these variables and add to the evidence base underlying efforts to improve the development of self-regulatory skills and school readiness in low-income minority children.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by grants from the National Institute of Child Health and Human Development.

References

- Bassett HH, Denham S, Wyatt TM, & Warren-Khot HK (2012). Refining the Preschool Self-Regulation Assessment for use in preschool classrooms. *Infant and Child Development*, 21, 596–616. 10.1002/icd.1763
- Best JR, & Miller PH (2010). A developmental perspective on executive function. *Child Development*, 81, 1641–1660. 10.1111/j.1467-8624.2010.01499.x [PubMed: 21077853]

- Bierman KL, Nix R, Greenberg MT, Blair CB, & Domitrovich CE (2008). Executive functions and school readiness intervention: Impact, moderation, and mediation in the Head Start REDI program. *Development and Psychopathology*, 20, 821–843. 10.1017/s0954579408000394 [PubMed: 18606033]
- Blair CB, & Razza RP (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, 78, 647–663. 10.1111/j.1467-8624.2007.01019.x [PubMed: 17381795]
- Blair CB, & Ursache A (2011). A bidirectional model of executive functions and self-regulation In Vohs KD & Baumeister RF (Eds.), *Handbook of self-regulation: Research, theory, and applications* (2nd ed., pp. 300–320). New York, NY: Guilford.
- Blair CB, Ursache A, Greenberg M, & Vernon-Feagans L (2015). Multiple aspects of self-regulation uniquely predict mathematics but not letter–word knowledge in the early elementary grades. *Developmental Psychology*, 51, 459–472. 10.1037/a0038813 [PubMed: 25688999]
- Brock LL, Rimm-Kaufman SE, Nathanson L, & Grimm KJ (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, 337–349. 10.1016/j.ecresq.2009.06.001
- Bull R, Espy KA, & Wiebe SA (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, 33, 205–228. 10.1080/87565640801982312 [PubMed: 18473197]
- Caughy MO, Owen MT, Mills BA, & Hurst JR (2013). Emergent self-regulation skills among very young ethnic minority children: A confirmatory factor model. *Journal of Experimental Child Psychology*, 116(4), 839–855. 10.1016/j.jecp.2013.07.017. [PubMed: 24076382]
- Cepeda NJ, Blackwell KA, & Munakata Y (2013). Speed isn’t everything: Complex processing speed measures mask individual differences and developmental changes in executive control. *Developmental Science*, 16, 269–286. 10.1111/desc.12024 [PubMed: 23432836]
- Cronbach LJ, & Furby L (1970). How we should measure” change”: Or should we? *Psychological Bulletin*, 74,68 10.1037/h0029382
- Denham SA, Warren-Khot HK, Bassett HH, Wyatt T, & Perna A (2012). Factor structure of selfregulation in preschoolers: Testing models of a fieldbased assessment for predicting early school readiness. *Journal of Experimental Child Psychology*, 111, 386–404. 10.1016/j.jecp.2011.10.002 [PubMed: 22104321]
- Dennis M, Francis DJ, Cirino PT, Schachar R, Barnes MA, & Fletcher JM (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society*, 15, 331–343. 10.1017/S1355617709090481 [PubMed: 19402919]
- Diamond A, Carlson SM, & Beck DM (2005). Preschool children’s performance in task switching on the Dimensional Change Card Sort task: Separating the dimensions aids the ability to switch. *Developmental Neuropsychology*, 28, 689–729. 10.1207/s15326942dn2802_7 [PubMed: 16144433]
- Diamond A, & Lee K (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, 333, 959–964. 10.1126/science.1204529 [PubMed: 21852486]
- Espy KA, McDiarmid MM, Cwik MF, Stalets MM, Hamby A, & Senn TE (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology*, 26, 465–486. 10.1207/s15326942dn2601_6 [PubMed: 15276905]
- Ferrer E, & McArdle JJ (2004). An experimental analysis of dynamic hypotheses about cognitive abilities and achievement from childhood to early adulthood. *Developmental Psychology*, 40, 935 10.1037/0012-1649.40.6.935 [PubMed: 15535749]
- Friedman NP, Miyake A, Robinson JL, & Hewitt JK (2011). Developmental trajectories in toddlers’ selfrestraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology*, 47, 1410 <https://doi.org/10.1037%2Fa0023750> [PubMed: 21668099]

- Fuhs MW, Nesbitt KT, Farran DC, & Dong N (2014). Longitudinal associations between executive functioning and academic skills across content areas. *Developmental Psychology*, 50, 1698–1709. 10.1037/a0036633 [PubMed: 24749550]
- Gardner MF (1985). *Receptive One-Word Picture Vocabulary Test*. Novato, CA: Academic Therapy.
- Garon N, Bryson SE, & Smith IM (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, 134, 32–60. 10.1037/0033-2909.134.1.31
- Grice JW (2001). Computing and evaluating factor scores. *Psychological Methods*, 6, 430. 10.1037/1082-989X.6.4.430 [PubMed: 11778682]
- Hughes C, Ensor R, Wilson A, & Graham A (2010). Tracking executive function across the transition to school: A latent variable approach. *Developmental Neuropsychology*, 35, 20–36. 10.1080/87565640903325691 [PubMed: 20390590]
- Kochanska G, Coy KC, & Murray K (2001). The development of self-regulation in the first four years of life. *Child Development*, 72, 1091–1111. 10.1111/1467-8624.00336 [PubMed: 11480936]
- Kochanska G, Murray K, & Harlan ET (2000). Effortful control in early childhood: Continuity and change, antecedents and implications for social development. *Developmental Psychology*, 36, 220–232. 10.1037/0012-1649.36.2.220 [PubMed: 10749079]
- Lan X, Legare CH, Ponitz CC, Li S, & Morrison FJ (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, 677–692. 10.1016/j.jecp.2010.11.001 [PubMed: 21238977]
- Liew J (2012). Effortful control, executive functions, and education: Bringing self-regulatory and social-emotional competencies to the table. *Child Development Perspectives*, 6, 105–111. 10.1111/j.1750-8606.2011.00196.x
- Li-Grining CP, Votruba-Drzal E, Maldonado-Carreño C, & Haas K (2010). Children's early approaches to learning and academic trajectories through fifth grade. *Developmental Psychology*, 46, 1062–1077. 10.1037/a0020066 [PubMed: 20822223]
- Lord FM (1967). A paradox in the interpretation of group comparisons. *Psychological Bulletin*, 68, 304–305. 10.1037/h0025105 [PubMed: 6062585]
- Matthews JS, Kizzie KT, Rowley SJ, & 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, 757–771. 10.1037/a0019616
- Matthews JS, Ponitz CC, & Morrison FJ (2009). Early gender differences in self-regulation and academic achievement. *Journal of Educational Psychology*, 101, 689–704. 10.1037/a0014240
- McArdle JJ (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, 60, 577–605. 10.1146/annurev.psych.60.110707.163612
- McClelland MM, Acock AC, & Morrison FJ (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly*, 21, 471–490. 10.1016/j.ecresq.2006.09.003
- McClelland MM, Cameron CE, Connor CM, Farris CL, Jewkes AM, & Morrison FJ (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*, 43, 947–959. 10.1037/0012-1649.43.4.947 [PubMed: 17605527]
- McClelland MM, Cameron CE, Duncan R, Bowles RP, Acock AC, Miao A, & Pratt ME (2014). Predictors of early growth in academic achievement: The Head-Toes-Knees-Shoulders task. *Frontiers in Psychology*, 5, 599. 10.3389/fpsyg.2014.00599 [PubMed: 25071619]
- McClelland MM, Geldhof GJ, Cameron CE, & Wanless SB (2015). Development and self-regulation. In Lerner RM (Ed.), *Handbook of child psychology and developmental science* (Vol. 1, 7th ed., pp. 523–565). New York, NY: Wiley.
- Miller MR, Giesbrecht GF, Müller U, McInerney RJ, & Kerns KA (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development*, 13, 395–423. 10.1080/15248372.2011.585478
- Mischel W, Shoda Y, & Rodriguez ML (1989). Delay of gratification in children. *Science*, 244, 933–938. 10.1126/science.2658056 [PubMed: 2658056]
- Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, & Wager TD (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent

- variable analysis. *Cognitive Psychology*, 41, 49–100. 10.1006/cogp.1999.0734 [PubMed: 10945922]
- Muthén B, & Muthén L (1998–2006). *Mplus user's guide* (4th ed.). Los Angeles, CA: Author.
- NICHD ECCRN. (1998). Early child care and self-control, compliance and problem behavior at 24 and 36 months. *Child Development*, 69, 1145–1170. 10.2307/1132367 [PubMed: 9768491]
- Petersen IT, Hoyniak CP, McQuillan ME, Bates JE, & Staples AD (2016). Measuring the development of inhibitory control: The challenge of heterotypic continuity. *Developmental Review*, 40, 25–71. 10.1016/j.dr.2016.02.001 [PubMed: 27346906]
- Ponitz CC, McClelland MM, Jewkes AM, Connor CM, Farris CL, & Morrison FJ (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly*, 23, 141–158. 10.1016/j.ecresq.2007.01.004
- Ponitz CC, McClelland MM, Matthews JS, & Morrison FJ (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology*, 45, 605–619. 10.1037/a0015365 [PubMed: 19413419]
- Posner M, & Rothbart M (1992). Attentional mechanisms and conscious experience In Milner AD & Rugg MD (Eds.), *Neuropsychology of consciousness* (pp. 91–111). London, UK: Academic Press 10.1016/B978-0-12-498045-7.50010-4
- Raver CC (2004). Placing emotional self-regulation in sociocultural and socioeconomic contexts. *Child Development*, 75, 346–353. 10.1111/j.1467-8624.2004.00676.x [PubMed: 15056189]
- Raver CC, Jones SM, Li-Grining C, Zhai F, Bub K, & Pressler E (2011). CSRP's impact on low-income preschoolers' preacademic skills: Self-regulation as a mediating mechanism. *Child Development*, 82, 362–378. 10.1111/j.1467-8624.2010.01561.x [PubMed: 21291447]
- Rhoades BL, Greenberg MT, Lanza ST, & Blair CB (2011). Demographic and familial predictors of early executive function development: Contribution of a person-centered perspective. *Journal of Experimental Child Psychology*, 108, 638–662. 10.1016/j.jecp.2010.08.004 [PubMed: 20828709]
- Schmitt SA, Geldof GJ, Purpura DJ, Duncan R, & McClelland MM (2017). Examining the relations between executive function, math, and literacy during the transition to kindergarten: A multi-analytic approach. *Journal of Educational Psychology*, 109, 1120 10.1037/edu0000193
- Sektan M, McClelland MM, Acock A, & Morrison FJ (2010). Relations between early family risk, children's behavioral regulation and academic achievement. *Early Childhood Research Quarterly*, 25, 464–479. 10.1016/j.ecresq.2010.02.005 [PubMed: 20953343]
- Selig JP, & Preacher KJ (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development*, 6, 144–164. 10.1080/15427600902911247
- Steenkamp JBE, & Baumgartner H (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25, 78–107. 10.1086/209528
- Ursache A, Blair CB, & Raver CC (2012). The promotion of self-regulation as a means of enhancing school readiness and early achievement in children at risk for school failure. *Child Development Perspectives*, 6, 122–128. 10.1111/j.1750-8606.2011.00209.x
- U.S. Department of Education. (2004). From kindergarten through third grade: Children's beginning school experiences. Washington, DC: National Center for Education Statistics Retrieved from http://nces.ed.gov/programs/quarterly/vol_6/6_3/2_1.asp
- Vandenberg RJ, & Lance CE (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4–70. 10.1177/109442810031002
- Welsh JA, Nix RL, Blair CB, Bierman KL, & Nelson KE (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology*, 102, 43–53. 10.1037/a0016738 [PubMed: 20411025]
- Widaman KF, Ferrer E, & Conger RD (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives*, 4, 10–18. 10.1111/j.1750-8606.2009.00110.x [PubMed: 20369028]
- Wiebe SA, Espy KA, & Charak D (2008). Using con-firmatory factor analysis to understand executive control in preschool children: I. Latent Structure. *Developmental Psychology*, 44, 575–587. 10.1037/0012-1649.44.2.575 [PubMed: 18331145]

- Wiebe SA, Sheffield TD, & Espy KA (2012). Separating the fish from the sharks: A longitudinal study of preschool response inhibition. *Child Development*, 83, 1245–1261. 10.1111/j.1467-8624.2012.01765.x [PubMed: 22533453]
- Wiebe SA, Sheffield T, Nelson JM, Clark CAC, Chevalier N, & Espy KA (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology*, 108, 436–452. 10.1016/j.jecp.2010.08.008 [PubMed: 20884004]
- Willoughby MT, Blair CB, Wirth RJ, & Greenberg M; The Family Life Project Investigators. (2010). The measurement of executive function at age 3 years: Psychometric properties and criterion validity of a new battery of tasks. *Psychological Assessment*, 22, 306–317. 10.1037/a0018708 [PubMed: 20528058]
- Willoughby MT, Kupersmidt JB, & Voegler-Lee ME (2012). Is preschool executive function causally related to academic achievement? *Child Neuropsychology*, 18, 79–91. 10.1080/09297049.2011.578572 [PubMed: 21707258]
- Willoughby MT, Pek J, & Blair CB (2013). Measuring executive function in early childhood: A focus on maximal reliability and the derivation of short forms. *Psychological Assessment*, 25, 664–670. 10.1037/a0031747 [PubMed: 23397928]
- Willoughby MT, Wirth RJ, & Blair CB (2012). Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological Assessment*, 24, 418–431. 10.1037/a0025779 [PubMed: 22023561]
- Woodcock RW (1990). Theoretical foundations of the WJ–R measures of cognitive ability. *Journal of Psychoeducational Assessment*, 8, 231–258. 10.1177/073428299000800303
- Woodcock RW, & Munoz-Sandoval AF (1996). *Bateria Woodcock-Munoz: Pruebas de aprovechamiento-Revisada* (Woodcock-Munoz Battery: Revised) Itasca, IL: Riverside.

Figure 1.

Cross-lagged (a) and latent difference score (LDS) (b) parameterizations of a longitudinal model of transactional relations across time. X , Y , and Z are hypothetical latent (X and Y) and manifest (Z) variables. For simplicity, manifest indicators of latent variables are not depicted. For a given variable X , LDS parameterization constraints (shown in gray) partition $X(t)$ into two components—the variable's value at the previous time point $X(t-1)$ and the current deviation from that value $\Delta X(t)$ —by fixing each component's effect on $X(t)$ to 1 and constraining the residual variance of $X(t)$ to 0 (not shown).

Figure 2.

Estimated latent variables and the parameterization of change across the four study time points. Latent difference score parameterization constraints are depicted in gray. Asterisks identify expanded measures that retained items from previous assessment points. Correlations, couplings, and cross-lagged effects that were estimated in the models are not shown.

Table 1

Study Sample Characteristics (N = 403)

	African American (N = 182) N (%)	Latino (N = 221) N (%)	χ^2
Child gender			
Boy	100 (54.9)	116 (52.5)	0.24
Girl	82 (45.1)	105 (47.5)	
Primary caregiver's relationship to child			
Mother	161 (88.5)	215 (97.3)	12.44 **
Father	11 (6.0)	3 (1.4)	
Grandmother/aunt	10 (5.5)	3 (1.4)	
Primary caregiver's ethnicity			
Non-Latino	182 (100.0)	12 (5.4)	357.54 ***
Latino	0 (0.0)	209 (94.6)	
Average family income-to-needs ratio			
< 50% federal poverty level	88 (49.7)	24 (10.9)	75.73 ***
50%–99% federal poverty level	53 (29.9)	107 (48.4)	
100%–149% federal poverty level	19 (10.7)	60 (27.1)	
150%+ federal poverty level	17 (9.6)	30 (13.6)	
Primary caregiver's level of education			
Less than high school	27 (14.8)	94 (42.5)	38.21 ***
High school/general education diploma	83 (45.6)	78 (35.3)	
More than high school	72 (39.6)	49 (22.2)	
Father/father-figure in household			
No	74 (40.7)	22 (10.0)	51.85 ***
Yes	108 (59.3)	199 (90.0)	
Latino families only			
Caregiver language status			
Spanish-dominant		125 (56.6)	
English-dominant		11 (5.0)	
Bilingual		71 (32.1)	
Missing		14 (6.3)	
Preferred language, Wave 1 (child)			
Spanish		174 (78.7)	
English		47 (21.3)	
Nativity (caregiver)			
U.S. born		60 (27.1)	
Foreign born		161 (72.9)	
Country of origin (foreign-born caregivers)			
Mexico		154 (95.7)	
Central America		6 (3.7)	

	African American (<i>N</i> = 182) <i>N</i> (%)	Latino (<i>N</i> = 221) <i>N</i> (%)	χ^2
South America		1 (.6)	

p < .01.

p < .001.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Key Parameter Estimates From Models of Transactional Relations Between Self-Regulation Measures and Subsequent Academic Achievement

Table 2

Variable	Mean/intercept	Proportional change	Variance	Prospective transactional effects on endogenous outcomes							
				Self-regulation (Wave 3)		Reading (Waves 3–4)		Math (Waves 3–4)			
				CI	SS	OS	WJR	WJR	WJM	WJM	WJM
FS	0	—	1	.37 (.17)*	-.16 (.18)	-.30 (.19)	.18 (.25)	.19 (.27)	.19 (.12)	.12 (.13)	.12 (.13)
SI	0	-.90 (.09)***	1	.23 (.17)	.77 (.29)**	.56 (.21)**	.22 (.22)	-.43 (.33)	.32 (.10)	-.20 (.13)	-.20 (.13)
SI	1.79 (.24)***	—	.87 (.25)***	.11 (.11)	.48 (.21)*	.24 (.15)	.23 (.17)	-.32 (.22)	.26 (.07)	-.10 (.08)	-.10 (.08)
CI	0	-.95 (.17)***	1	—	-.22 (.15)	-.25 (.17)	.22 (.23)	-.04 (.21)	.14 (.09)	.15 (.08)	.15 (.08)
CI	1.91 (.34)***	—	1.52 (.44)***	—	—	—	—	-.09 (.12)	—	.00 (.05)	.00 (.05)
SS	0	-.86 (.12)***	1	-.07 (.12)	—	-.03 (.14)	.19 (.18)	.37 (.22)	-.06 (.09)	.12 (.08)	.12 (.08)
SS	1.52 (.41)***	—	.84 (.43)*	—	—	—	—	.49 (.19)**	—	.14 (.08)	.14 (.08)
Memory chocolates (MC)	0	—	1	.16 (.13)	.33 (.16)*	.26 (.15)	.16 (.16)	-.12 (.16)	.05 (.09)	-.05 (.07)	-.05 (.07)
Operation span (OS)	0	—	1	—	—	—	—	-.18 (.20)	—	.12 (.08)	.12 (.08)
WJ reading (WJR)	-4.88 (.42)***	-.39 (.06)***	3.63 (.34)***	—	—	—	—	—	—	—	—
WJR	4.60 (.70)***	—	1.51 (.17)***	—	—	—	—	—	—	—	—
WJ math (WJM)	-3.95 (.16)***	-.44 (.06)***	.70 (.06)***	—	—	—	—	—	—	—	—
WJM	3.26 (.34)***	—	.28 (.04)***	—	—	—	—	—	—	—	—

Note. Standard errors are shown in parentheses. Fixed parameter values are shown with no standard error. Dashes indicate parameters not estimated because they are nonsensical (e.g., they implied reverse causality; variables not parameterized as latent differences do not have proportional change parameters). WJM estimates were obtained from a separate model run using WJM and WJM as outcomes instead of WJR and WJR; model specifications were otherwise identical. Estimates common to reading and math models (168 of 195 estimated parameters) were virtually identical and differed in no substantive way. Global fit indices were acceptable and identical for both models (RMSEA = .02, CFI = .97, TLI (Tucker-Lewis Index) = .96). Variables measured at each wave were FS and SI (Wave 1); SI, CI, SS, and MC (Wave 2); CI, DSS, OS, WJR, and WJM (Wave 3); and WJR and WJM (Wave 4). Nonsignificant prospective effects of the Wave 1 variables on Wave 2 variables are not shown. Residual correlations between variables at the same time point are provided in Table S2. All effects on WJ outcomes adjust for participant ethnicity and family income-to-needs ratio. WJ = Woodcock-Johnson; FS = fruit stroop; SI = simple response inhibition; CI = complex response inhibition; SS = set shifting; RMSEA = root mean square error of approximation; CFI = comparative fit index.

* $p < .05$.

** $p < .01$.

*** $p < .001$.